

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OW protein - protein search, using sw model

Run on: December 30, 2004, 16:04:46 ; Search time 37 Seconds
(without alignments)
32.263 Million cell updates/sec

Title: US-10-718-321-1
Perfect score: 99
Sequence: 1 SSDGLMNNQTLFLEHS 18

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 150 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/PTCUS_COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	99	100.0	334	4	US-09-197-9708-7
2	84	84.8	451	1	US-08-287-001A-2
3	84	84.8	451	5	PCT-US95-09941-2
4	52	52.5	570	4	US-09-248-796A-15079
5	48	48.5	981	4	US-09-991-258-13
6	45	45.5	114	6	5514582-13
7	45	45.5	114	6	5514582-14
8	44	44.4	1464	1	US-08-026-138E-1
9	44	44.4	590	4	US-09-538-092-967
10	43	43.4	62	4	US-09-248-796A-22793
11	43	43.4	251	4	US-09-270-767-9147
12	43	43.4	265	4	US-09-270-767-47089
13	43	43.4	311	4	US-09-248-796A-20325
14	42	42.4	187	4	US-09-540-236-3758
15	42	42.4	358	4	US-09-252-991A-21474
16	42	42.4	1190	4	US-09-248-796A-26719
17	41	41.4	120	4	US-09-248-796A-14464
18	41	41.4	155	4	US-09-134-000C-4868
19	41	41.4	166	4	US-08-089-397A-11
20	41	41.4	326	6	5395759-2
21	41	41.4	365	4	US-09-248-796A-18632
22	41	41.4	817	4	US-09-248-796A-17089
23	41	41.4	1455	3	US-08-840-062-5
24	41	41.4	1464	2	US-08-231-193A-11
25	41	41.4	1464	2	US-08-466-273A-11
26	41	41.4	1464	2	US-08-940-086A-11
27	41	41.4	1464	3	US-08-436-332B-10
28	41	41.4	1464	3	US-08-940-035A-11
29	41	41.4	1464	3	US-08-935-105A-11
30	41	41.4	1464	4	US-09-648-797-11
31	41	41.4	1464	4	US-08-217-704C-2
32	41	41.4	1464	4	US-09-386-123-11
33	41	41.4	1464	4	US-09-187-049-8
34	40	40.4	126	3	US-09-540-326-3527
35	40	40.4	141	4	US-09-513-999C-5695
36	40	40.4	186	4	US-09-248-796A-16186
37	40	40.4	421	2	US-08-807-263-4
38	40	40.4	463	4	US-09-375-140-10
39	40	40.4	489	2	US-08-895-521-1
40	40	40.4	489	2	US-08-895-521-3
41	40	40.4	489	2	US-09-235-218-1
42	40	40.4	489	3	US-09-235-218-3
43	40	40.4	495	4	US-08-899-324-33
44	40	40.4	532	2	US-08-899-324-33
45	40	40.4	532	3	US-08-329-892B-33
46	40	40.4	536	3	US-08-653-648A-13
47	40	40.4	536	4	US-09-564-418-9
48	40	40.4	548	4	US-09-375-140-11
49	40	40.4	571	1	US-08-419-009-1
50	40	40.4	803	4	US-09-543-681A-4886
51	40	40.4	830	4	US-09-562-737-34
52	40	40.4	1198	3	US-09-199-637A-405
53	39.5	39.9	122	4	US-09-800-729-147
54	39	39.4	122	4	US-09-540-326-3223
55	39	39.4	130	4	US-09-538-092-284
56	39	39.4	130	4	US-09-691-270A-24
57	39	39.4	543	3	US-09-155-768-4
58	39	39.4	557	1	US-08-309-341-2
59	39	39.4	557	1	US-08-608-267-2
60	39	39.4	557	1	US-08-608-267-4
61	39	39.4	557	1	US-08-608-452-2
62	39	39.4	557	1	US-08-608-452-4
63	39	39.4	557	1	US-08-608-452-4
64	39	39.4	557	1	US-08-608-452-4
65	39	39.4	557	1	US-08-608-452-4
66	39	39.4	557	1	US-08-608-452-4
67	39	39.4	557	1	US-08-608-452-4
68	39	39.4	557	1	US-08-608-452-4
69	39	39.4	557	1	US-08-608-452-4
70	39	39.4	557	1	US-08-608-452-4
71	39	39.4	557	1	US-08-608-452-4
72	39	39.4	557	1	US-08-608-452-4
73	39	39.4	557	1	US-08-608-452-4
74	39	39.4	557	1	US-08-608-452-4
75	39	39.4	557	1	US-08-608-452-4
76	39	39.4	557	1	US-08-608-452-4
77	38.5	38.9	429	4	US-09-252-991A-31588
78	38.5	38.9	652	6	5258288-4
79	38	38.4	21	3	US-09-252-586-18
80	38	38.4	93	4	US-09-248-796A-21642
81	38	38.4	129	5	PCT-US95-04258-10
82	38	38.4	129	5	PCT-US95-04258-10
83	38	38.4	137	4	US-09-270-767-75914
84	38	38.4	150	4	US-09-270-767-35523
85	38	38.4	150	4	US-09-270-767-50740
86	38	38.4	155	3	US-08-608-452-4
87	38	38.4	155	3	US-08-608-452-4
88	38	38.4	155	3	US-08-608-452-4
89	38	38.4	155	3	US-08-608-452-4
90	38	38.4	155	3	US-08-608-452-4
91	38	38.4	155	3	US-08-608-452-4
92	38	38.4	155	3	US-08-608-452-4
93	38	38.4	155	3	US-08-608-452-4
94	38	38.4	155	3	US-08-608-452-4
95	38	38.4	155	3	US-08-608-452-4
96	38	38.4	155	3	US-08-608-452-4
97	38	38.4	155	3	US-08-608-452-4
98	38	38.4	155	3	US-08-608-452-4
99	38	38.4	155	3	US-08-608-452-4
100	38	38.4	155	3	US-08-608-452-4

101	38	38.4	535	3	US-08-286-870A-6	Sequence 6, Appli
102	38	38.4 <td>543<td>2</td><td>US-08-922-170B-10</td><td>Sequence 10, Appli</td></td>	543 <td>2</td> <td>US-08-922-170B-10</td> <td>Sequence 10, Appli</td>	2	US-08-922-170B-10	Sequence 10, Appli
103	38	38.4 <td>543<td>3</td><td>US-09-071-739B-2</td><td>Sequence 2, Appli</td></td>	543 <td>3</td> <td>US-09-071-739B-2</td> <td>Sequence 2, Appli</td>	3	US-09-071-739B-2	Sequence 2, Appli
104	38	38.4 <td>543<td>3</td><td>US-09-181-336-13</td><td>Sequence 13, Appli</td></td>	543 <td>3</td> <td>US-09-181-336-13</td> <td>Sequence 13, Appli</td>	3	US-09-181-336-13	Sequence 13, Appli
105	38	38.4 <td>543<td>3</td><td>US-09-260-038B-2</td><td>Sequence 2, Appli</td></td>	543 <td>3</td> <td>US-09-260-038B-2</td> <td>Sequence 2, Appli</td>	3	US-09-260-038B-2	Sequence 2, Appli
106	38	38.4 <td>543<td>4</td><td>US-09-635-923-2</td><td>Sequence 2, Appli</td></td>	543 <td>4</td> <td>US-09-635-923-2</td> <td>Sequence 2, Appli</td>	4	US-09-635-923-2	Sequence 2, Appli
107	38	38.4 <td>543<td>4</td><td>US-09-487-716A-2</td><td>Sequence 2, Appli</td></td>	543 <td>4</td> <td>US-09-487-716A-2</td> <td>Sequence 2, Appli</td>	4	US-09-487-716A-2	Sequence 2, Appli
108	38	38.4 <td>543<td>4</td><td>US-09-322-977-2</td><td>Sequence 2, Appli</td></td>	543 <td>4</td> <td>US-09-322-977-2</td> <td>Sequence 2, Appli</td>	4	US-09-322-977-2	Sequence 2, Appli
109	38	38.4 <td>543<td>4</td><td>US-09-186-200-1</td><td>Sequence 1, Appli</td></td>	543 <td>4</td> <td>US-09-186-200-1</td> <td>Sequence 1, Appli</td>	4	US-09-186-200-1	Sequence 1, Appli
110	38	38.4 <td>543<td>4</td><td>US-09-435-739-10</td><td>Sequence 10, Appli</td></td>	543 <td>4</td> <td>US-09-435-739-10</td> <td>Sequence 10, Appli</td>	4	US-09-435-739-10	Sequence 10, Appli
111	38	38.4 <td>543<td>4</td><td>US-09-930-218-3</td><td>Sequence 3, Appli</td></td>	543 <td>4</td> <td>US-09-930-218-3</td> <td>Sequence 3, Appli</td>	4	US-09-930-218-3	Sequence 3, Appli
112	38	38.4 <td>543<td>4</td><td>US-09-704-772A-2</td><td>Sequence 2, Appli</td></td>	543 <td>4</td> <td>US-09-704-772A-2</td> <td>Sequence 2, Appli</td>	4	US-09-704-772A-2	Sequence 2, Appli
113	38	38.4 <td>543<td>4</td><td>US-09-988-113-10</td><td>Sequence 10, Appli</td></td>	543 <td>4</td> <td>US-09-988-113-10</td> <td>Sequence 10, Appli</td>	4	US-09-988-113-10	Sequence 10, Appli
114	38	38.4 <td>545<td>4</td><td>US-09-899-440-18</td><td>Sequence 10, Appli</td></td>	545 <td>4</td> <td>US-09-899-440-18</td> <td>Sequence 10, Appli</td>	4	US-09-899-440-18	Sequence 10, Appli
115	38	38.4 <td>565<td>4</td><td>US-09-328-352-4920</td><td>Sequence 4920, Ap</td></td>	565 <td>4</td> <td>US-09-328-352-4920</td> <td>Sequence 4920, Ap</td>	4	US-09-328-352-4920	Sequence 4920, Ap
116	38	38.4 <td>587<td>2</td><td>US-08-642-684-2</td><td>Sequence 2, Appli</td></td>	587 <td>2</td> <td>US-08-642-684-2</td> <td>Sequence 2, Appli</td>	2	US-08-642-684-2	Sequence 2, Appli
117	38	38.4 <td>587<td>3</td><td>US-09-157-397-2</td><td>Sequence 2, Appli</td></td>	587 <td>3</td> <td>US-09-157-397-2</td> <td>Sequence 2, Appli</td>	3	US-09-157-397-2	Sequence 2, Appli
118	38	38.4 <td>587<td>3</td><td>US-09-517-871-22</td><td>Sequence 22, Appli</td></td>	587 <td>3</td> <td>US-09-517-871-22</td> <td>Sequence 22, Appli</td>	3	US-09-517-871-22	Sequence 22, Appli
119	38	38.4 <td>587<td>4</td><td>US-09-512-021-2</td><td>Sequence 2, Appli</td></td>	587 <td>4</td> <td>US-09-512-021-2</td> <td>Sequence 2, Appli</td>	4	US-09-512-021-2	Sequence 2, Appli
120	38	38.4 <td>588<td>3</td><td>US-09-157-397-4</td><td>Sequence 4, Appli</td></td>	588 <td>3</td> <td>US-09-157-397-4</td> <td>Sequence 4, Appli</td>	3	US-09-157-397-4	Sequence 4, Appli
121	38	38.4 <td>588<td>4</td><td>US-09-601-777-2</td><td>Sequence 2, Appli</td></td>	588 <td>4</td> <td>US-09-601-777-2</td> <td>Sequence 2, Appli</td>	4	US-09-601-777-2	Sequence 2, Appli
122	38	38.4 <td>588<td>4</td><td>US-09-512-021-4</td><td>Sequence 4, Appli</td></td>	588 <td>4</td> <td>US-09-512-021-4</td> <td>Sequence 4, Appli</td>	4	US-09-512-021-4	Sequence 4, Appli
123	38	38.4 <td>592<td>2</td><td>US-08-642-684-4</td><td>Sequence 4, Appli</td></td>	592 <td>2</td> <td>US-08-642-684-4</td> <td>Sequence 4, Appli</td>	2	US-08-642-684-4	Sequence 4, Appli
124	38	38.4 <td>592<td>4</td><td>US-09-435-739-14</td><td>Sequence 14, Appli</td></td>	592 <td>4</td> <td>US-09-435-739-14</td> <td>Sequence 14, Appli</td>	4	US-09-435-739-14	Sequence 14, Appli
125	38	38.4 <td>592<td>4</td><td>US-09-988-113-14</td><td>Sequence 14, Appli</td></td>	592 <td>4</td> <td>US-09-988-113-14</td> <td>Sequence 14, Appli</td>	4	US-09-988-113-14	Sequence 14, Appli
126	38	38.4 <td>621<td>4</td><td>US-09-071-035-468</td><td>Sequence 468, App</td></td>	621 <td>4</td> <td>US-09-071-035-468</td> <td>Sequence 468, App</td>	4	US-09-071-035-468	Sequence 468, App
127	38	38.4 <td>648<td>3</td><td>US-08-286-870A-4</td><td>Sequence 468, App</td></td>	648 <td>3</td> <td>US-08-286-870A-4</td> <td>Sequence 468, App</td>	3	US-08-286-870A-4	Sequence 468, App
128	38	38.4 <td>655<td>4</td><td>US-09-248-796A-14308</td><td>Sequence 14308, A</td></td>	655 <td>4</td> <td>US-09-248-796A-14308</td> <td>Sequence 14308, A</td>	4	US-09-248-796A-14308	Sequence 14308, A
129	38	38.4 <td>717<td>4</td><td>US-09-710-279-1022</td><td>Sequence 3022, Ap</td></td>	717 <td>4</td> <td>US-09-710-279-1022</td> <td>Sequence 3022, Ap</td>	4	US-09-710-279-1022	Sequence 3022, Ap
130	38	38.4 <td>719<td>3</td><td>US-08-286-870A-8</td><td>Sequence 8, Appli</td></td>	719 <td>3</td> <td>US-08-286-870A-8</td> <td>Sequence 8, Appli</td>	3	US-08-286-870A-8	Sequence 8, Appli
131	38	38.4 <td>765<td>3</td><td>US-09-134-001C-3681</td><td>Sequence 3681, Ap</td></td>	765 <td>3</td> <td>US-09-134-001C-3681</td> <td>Sequence 3681, Ap</td>	3	US-09-134-001C-3681	Sequence 3681, Ap
132	38	38.4 <td>849</td> <td>4</td> <td>US-09-248-796A-15566</td> <td>Sequence 15566, A</td>	849	4	US-09-248-796A-15566	Sequence 15566, A
133	38	38.4 <td>954<td>4</td><td>US-09-570-856B-31</td><td>Sequence 31, Appli</td></td>	954 <td>4</td> <td>US-09-570-856B-31</td> <td>Sequence 31, Appli</td>	4	US-09-570-856B-31	Sequence 31, Appli
134	38	38.4 <td>1402</td> <td>4</td> <td>US-09-248-796A-14503</td> <td>Sequence 14503, A</td>	1402	4	US-09-248-796A-14503	Sequence 14503, A
135	38	38.4 <td>2032</td> <td>4</td> <td>US-09-071-035-458</td> <td>Sequence 458, App</td>	2032	4	US-09-071-035-458	Sequence 458, App
136	38	38.4 <td>2032</td> <td>4</td> <td>US-09-071-035-462</td> <td>Sequence 462, App</td>	2032	4	US-09-071-035-462	Sequence 462, App
137	38	38.4 <td>2032</td> <td>4</td> <td>US-09-071-035-466</td> <td>Sequence 466, App</td>	2032	4	US-09-071-035-466	Sequence 466, App
138	38	38.4 <td>2054</td> <td>4</td> <td>US-09-134-000C-6612</td> <td>Sequence 6612, Ap</td>	2054	4	US-09-134-000C-6612	Sequence 6612, Ap
139	38	38.4 <td>2186</td> <td>2</td> <td>US-08-822-445-2</td> <td>Sequence 2, Appli</td>	2186	2	US-08-822-445-2	Sequence 2, Appli
140	38	38.4 <td>2186</td> <td>3</td> <td>US-09-396-540-2</td> <td>Sequence 1, Appli</td>	2186	3	US-09-396-540-2	Sequence 1, Appli
141	37.5	37.9	200	3	US-09-206-936-1	Sequence 30, Appli
142	37.5	37.9	262	3	US-08-961-083-30	Sequence 30, Appli
143	37.5	37.9	262	4	US-09-536-784-30	Sequence 30, Appli
144	37	37.4 <td>84</td> <td>4</td> <td>US-09-599-632-4</td> <td>Sequence 4, Appli</td>	84	4	US-09-599-632-4	Sequence 4, Appli
145	37	37.4 <td>223</td> <td>4</td> <td>US-09-198-452A-448</td> <td>Sequence 448, App</td>	223	4	US-09-198-452A-448	Sequence 448, App
146	37	37.4 <td>244</td> <td>4</td> <td>US-09-543-681A-4528</td> <td>Sequence 4528, Ap</td>	244	4	US-09-543-681A-4528	Sequence 4528, Ap
147	37	37.4 <td>251</td> <td>3</td> <td>US-09-091-725-21</td> <td>Sequence 21, Appli</td>	251	3	US-09-091-725-21	Sequence 21, Appli
148	37	37.4 <td>263</td> <td>4</td> <td>US-09-691-270A-22</td> <td>Sequence 22, Appli</td>	263	4	US-09-691-270A-22	Sequence 22, Appli
149	37	37.4 <td>279</td> <td>4</td> <td>US-09-270-767-38329</td> <td>Sequence 38329, A</td>	279	4	US-09-270-767-38329	Sequence 38329, A
150	37	37.4 <td>279</td> <td>4</td> <td>US-09-270-767-53546</td> <td>Sequence 53546, A</td>	279	4	US-09-270-767-53546	Sequence 53546, A

ALIGNMENTS

RESULT 1
US-09-197-970B-7
Sequence 7, Application US/09197970B
Patent No. 6664385
GENERAL INFORMATION:
APPLICANT: Michele Sanicola-Nadel
Joseph V. Bonventre
Catherine A. Hession
Takaharu Ichimura
Henry Wei
Richard V. Cate
TITLE OF INVENTION: MODULATORS OF TISSUE REGENERATION
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Biogen, Inc.
STREET: 14 Cambridge Center
CITY: Cambridge
STATE: MA

COUNTRY: USA
ZIP: 02142
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09197,970B
FILING DATE: 23-No. 6664385-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/018,228
FILING DATE: 24-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Levine, Leslie M.
REGISTRATION NUMBER: 35,245
REFERENCE/DOCKET NUMBER: A010 PCT CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 679-2810
TELEFAX: (617) 679-2838
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 334 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-197-970B-7

Query Match 100.0%; Score 99; DB 4; Length 334;
Best Local Similarity 100.0%; Pred. No. 5.8e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 SSDGLMNNOTLFLHS 18
Db 264 SSDGLMNNOTLFLHS 281

RESULT 2
US-08-287-001A-2
Sequence 2, Application US/08287001A
Patent No. 5622861
GENERAL INFORMATION:
APPLICANT: KAPLAN, GERARDO
APPLICANT: FEINSTONE, STEPHEN M.
TITLE OF INVENTION: HEPATITIS A VIRUS RECEPTOR AND METHODS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: NEEDLE & ROSENBERG, P.C.
STREET: Suite 1200, The Candler Bldg, 127 Peachtree
STREET: Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/287,001A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Spitalt, Gwedolyn D.
REGISTRATION NUMBER: 36,016
REFERENCE/DOCKET NUMBER: 1414.621
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404/688-0770
TELEFAX: 404/688-9880

;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 451 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-287-001A-2

Query Match 84.8%; Score 84; DB 1; Length 451;
Best Local Similarity 88.9%; Pred. No. 2e-05;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SSDGLMNNQOTLPLEHS 18
|||
Db 344 SSDGLMNNQOTLPLEHS 361

RESULT 3

PCT-US95-09941-2
; Sequence 2, Application PC/TUS9509941
; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: HEPATITIS A VIRUS RECEPTOR AND METHODS

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NEEDLE & ROSENBERG, P.C.

; STREET: Suite 1200, The Candler Bldg, 127 Peachtree

; CITY: Atlanta

; STATE: Georgia

; COUNTRY: USA

; ZIP: 30303

; COMPUTER READABLE FORM:

; MEDIUM TYPE: floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/09941

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/287,001

; FILING DATE: 5 AUG 1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Spratt, Gwedoym D.

; REGISTRATION NUMBER: 36,016

; REFERENCE/DOCKET NUMBER: 1414.621

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 404/688-0770

; TELEFAX: 404/688-9880

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 451 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

PCT-US95-09941-2

Query Match 84.8%; Score 84; DB 5; Length 451;
Best Local Similarity 88.9%; Pred. No. 2e-05;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SSDGLMNNQOTLPLEHS 18
|||
Db 344 SSDGLMNNQOTLPLEHS 361

RESULT 4

US-09-248-796A-15079
; Sequence 15079, Application US/09248796A
; Patent No. 6747137

;; GENERAL INFORMATION:
;; APPLICANT: Keith Weinstein et al
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
;; FILE REFERENCE: 107196.132
;; CURRENT APPLICATION NUMBER: US/09/248,796A
;; CURRENT FILING DATE: 1999-02-12
;; PRIOR APPLICATION NUMBER: US 60/074,725
;; PRIOR FILING DATE: 1998-02-13
;; PRIOR APPLICATION NUMBER: US 60/096,409
;; PRIOR FILING DATE: 1998-08-13
;; NUMBER OF SEQ ID NOS: 28208
;; SEQ ID NO 15079
;; LENGTH: 570
;; TYPE: PRT
;; ORGANISM: Candida albicans
US-09-248-796A-15079

Query Match 52.5%; Score 52; DB 4; Length 570;
Best Local Similarity 57.1%; Pred. No. 3.2;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 2 SDGLMNNQOTLPL 15
|||
Db 2 NDGLMNNQOTLPL 15

RESULT 5

US-09-991-258-13
; Sequence 13, Application US/09991258
; Patent No. 6783939
; GENERAL INFORMATION:

; APPLICANT: Olmsted, Robert

; APPLICANT: Keith, Paula

; APPLICANT: Dryga, Sergey

; APPLICANT: Caley, Ian

; APPLICANT: Maughan, Maureen

; APPLICANT: Johnston, Robert

; APPLICANT: Davis, Nancy

; APPLICANT: Swanson, Ronald

; TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOMES WITH MODIFIED HIV GENES FOR USE

; FILE REFERENCE: 0113,0001U3

; CURRENT APPLICATION NUMBER: US/09/991,258

; CURRENT FILING DATE: 2001-11-16

; PRIOR APPLICATION NUMBER: 09/902,537

; PRIOR FILING DATE: 2001-07-09

; PRIOR APPLICATION NUMBER: 60/216,995

; PRIOR FILING DATE: 2000-07-07

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 13

; LENGTH: 981

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence; No. 6783939e =

US-09-991-258-13

Query Match 48.5%; Score 48; DB 4; Length 981;
Best Local Similarity 50.0%; Pred. No. 26;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 SSDGLMNNQOTLPL 16
|||
Db 489 SLDLMLNNQOMFWIQ 504

RESULT 6

5514582-13
; Patent No. 5514582
; APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.

```
/ TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
/ IMMUNOGLOBULINS
/ NUMBER OF SEQUENCES: 43
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/185,670
/ FILING DATE: 21-JAN-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 986,931
/ FILING DATE: 08-DEC-1992
/ APPLICATION NUMBER: 808,122
/ FILING DATE: 16-DEC-1991
/ APPLICATION NUMBER: 440,625
/ FILING DATE: 22-NOV-1989
/ APPLICATION NUMBER: 315,015
/ FILING DATE: 23-FEB-1989
/ SEQ ID NO:13:
/ LENGTH: 114
514582-13

Query Match          45.5%; Score 45; DB 6; Length 114;
Best Local Similarity 72.7%; Pred. No. 6.9;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 SDGLMNNQTO 12
Db      97 SDGLMNNFCQ 107

RESULT 7
514582-14
/ Patent No. 5514582
/ APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.
/ TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
/ IMMUNOGLOBULINS
/ NUMBER OF SEQUENCES: 43
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/185,670
/ FILING DATE: 21-JAN-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 986,931
/ FILING DATE: 08-DEC-1992
/ APPLICATION NUMBER: 808,122
/ FILING DATE: 16-DEC-1991
/ APPLICATION NUMBER: 440,625
/ FILING DATE: 22-NOV-1989
/ APPLICATION NUMBER: 315,015
/ FILING DATE: 23-FEB-1989
/ SEQ ID NO:14:
/ LENGTH: 114
514582-14

Query Match          45.5%; Score 45; DB 6; Length 114;
Best Local Similarity 72.7%; Pred. No. 6.9;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 SDGLMNNQTO 12
Db      97 SDGLMNNFCQ 107

RESULT 8
US-08-026-138E-1
/ Sequence 1, Application US/08026138E
/ Patent No. 5502166
/ GENERAL INFORMATION:
/ APPLICANT: MASAYOSHI MISHINA
/ TITLE OF INVENTION: NOVEL PROTEINS AND GENES CODING THE SAME
/ NUMBER OF SEQUENCES: 19
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Nishiohata Residence 1-107
/ STREET: 5214, Nishiohata-machi
/ CITY: Niigata-shi
/ STATE: Niigata-ken
```

```
/ COUNTRY: JAPAN
/ ZIP: 951
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: MS-DOS v.5
/ SOFTWARE: Word Perfect 5.1
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/026,138E
/ FILING DATE: 26-FEB-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: JP 39563/1992
/ FILING DATE: 26-FEB-1992
/ APPLICATION NUMBER: JP 173155/1992
/ FILING DATE: 30-JUN-1992
/ APPLICATION NUMBER: JP 215017/1992
/ FILING DATE: 12-AUG-1992
/ APPLICATION NUMBER: JP 303878/1992
/ FILING DATE: 13-NOV-1992
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Hamburg, C.Bruce
/ REGISTRATION NUMBER: 22,389
/ REFERENCE/DOCKET INFORMATION:
/ TELEPHONE: (212) 986-2340
/ TELEFAX: (212) 953-7733
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1464 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single strand
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ ORIGINAL SOURCE:
/ ORGANISM: mouse
/ TISSUE TYPE: cerebellum
/ PUBLICATION INFORMATION:
/ AUTHORS: Masayoshi MISHINA
/ TITLE: NOVEL PROTEINS AND GENES CODING THE SAME
/ RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 to 1464
US-08-026-138E-1

Query Match          45.5%; Score 45; DB 1; Length 1464;
Best Local Similarity 47.1%; Pred. No. 1.2e+02;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY      1 SSDGLMNNQTOLEH 17
Db      1174 NEDGLPNNQYKLYAKH 1190

RESULT 9
US-09-538-092-967
/ Sequence 967, Application US/09538092
/ Patent No. 6753514
/ GENERAL INFORMATION:
/ APPLICANT: Giot, Loic
/ TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
/ FILE REFERENCE: 15966-542
/ CURRENT APPLICATION NUMBER: US/09/538,092
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: 60/127,352
/ PRIOR FILING DATE: 1999-04-01
/ PRIOR APPLICATION NUMBER: 60/178,965
/ PRIOR FILING DATE: 2000-02-01
/ NUMBER OF SEQ ID NOS: 1387
/ SOFTWARE: CuratSeqFormatter Version 0.9
/ SEQ ID NO 967
/ LENGTH: 590
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
```

NAME/KEY: misc feature
LOCATION: (0)...(0)
OTHER INFORMATION: Polypeptide Accession Number P20309
US-09-538-092-967

Query Match 44.4%; Score 44; DB 4; Length 590;
Best Local Similarity 50.0%; Pred. No. 63;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 SSDGLMNNQTOFLFHS 18
DB 332 SSDSMNNNDAAASLENS 349

RESULT 10
US-09-248-796A-22793
Sequence 22793, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Kelch Weinstein et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 22793
LENGTH: 62
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-22793

Query Match 43.4%; Score 43; DB 4; Length 62;
Best Local Similarity 47.1%; Pred. No. 7.3;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 SDGLMNNQTOFLFHS 18
DB 25 SDSLQDTNFTYLFIOHS 41

RESULT 11
US-09-270-767-49147
Sequence 49147, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 49147
LENGTH: 251
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-49147

Query Match 43.4%; Score 43; DB 4; Length 251;
Best Local Similarity 56.2%; Pred. No. 35;
Matches 9; Conservative 4; Mismatches 1; Indels 2; Gaps 1;

QY 5 LMNN--QTOFLFHS 18
DB 52 LMNNYOKTRLFRKNS 67

RESULT 12

US-09-270-767-47089
Sequence 47089, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 47089
LENGTH: 265
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-47089

Query Match 43.4%; Score 43; DB 4; Length 265;
Best Local Similarity 56.2%; Pred. No. 37;
Matches 9; Conservative 4; Mismatches 1; Indels 2; Gaps 1;

QY 5 LMNN--QTOFLFHS 18
DB 66 LMNNYOKTRLFRKNS 81

RESULT 13
US-09-270-767-48661
Sequence 48661, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 48661
LENGTH: 311
TYPE: PRT
ORGANISM: Drosophila melanogaster
FEATURE:
OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-48661

Query Match 43.4%; Score 43; DB 4; Length 311;
Best Local Similarity 56.2%; Pred. No. 44;
Matches 9; Conservative 4; Mismatches 1; Indels 2; Gaps 1;

QY 5 LMNN--QTOFLFHS 18
DB 141 LMNNYOKTRLFRKNS 156

RESULT 14
US-09-248-796A-20325
Sequence 20325, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Kelch Weinstein et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 20325
LENGTH: 187

```
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-20325

Query Match          42.4%; Score 42; DB 4; Length 187;
Best Local Similarity 60.0%; Pred. No. 36;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 SSDGLMNNQTOFL 15
Db 9 SSDGLMNNKAKLPL 23

RESULT 15
US-09-540-236-3758
; Sequence 3758, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540.236
; NUMBER OF SEQ ID NOS: 2000-04-04
; SEQ ID NO 3758
; LENGTH: 358
; TYPE: PRT
; ORGANISM: M.cattarrhalis
US-09-540-236-3758

Query Match          42.4%; Score 42; DB 4; Length 358;
Best Local Similarity 63.6%; Pred. No. 75;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 GLMNNQTOFL 14
Db 80 GLMNDQEQAF 90

RESULT 16
US-09-252-991A-21474
; Sequence 21474, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074.788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094.190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21474
; LENGTH: 1190
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21474

Query Match          42.4%; Score 42; DB 4; Length 1190;
Best Local Similarity 50.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 SDGLMNNQTOFL 13
Db 346 TDGLMNNDSASY 357

RESULT 17
US-09-248-796A-26719
```

```
; Sequence 26719, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248.796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074.725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096.409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 26719
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-26719

Query Match          41.4%; Score 41; DB 4; Length 120;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 SDGLMNNQTOFL 13
Db 71 SDGLMNNLEEV 82

RESULT 18
US-09-248-796A-14464
; Sequence 14464, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248.796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074.725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096.409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 14464
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-14464

Query Match          41.4%; Score 41; DB 4; Length 155;
Best Local Similarity 46.7%; Pred. No. 42;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 DGLMNNQTOFLER 17
Db 29 DKWKVKNTQVFOEN 43

RESULT 19
US-09-134-000C-4868
; Sequence 4868, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134.000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055.778
; PRIOR FILING DATE: 1997-08-15
```

NUMBER OF SEQ ID NOS: 6812
SOFTWARE: Patentin version 3.1
SEQ ID NO 4868
LENGTH: 166
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-134-000C-4868

Query Match 41.4% Score 41; DB 4; Length 166;
Best Local Similarity 53.8% Pred. No. 46;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 SDGLMNNQTOLE 14
DB 86 SDRTFNNYTFIF 98

RESULT 20
US-08-089-397A-11
Sequence 11, Application US/08089397A
Patent No. 6086880
GENERAL INFORMATION:
APPLICANT: SABARA, MARTA I.J.
APPLICANT: FRENCHICK, PATRICK J.
APPLICANT: POTTER, ANDREW A.
APPLICANT: IJAZ, MOHAMMAD K.
APPLICANT: GILCHRIST, JAMES E.
APPLICANT: REDMOND, MARK J.
TITLE OF INVENTION: ROTAVIRUS VACCINES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1888
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/089.397A
FILING DATE: 07-JUL-1993
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 29311-20003.03
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 326 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-089-397A-11

Query Match 41.4% Score 41; DB 3; Length 326;
Best Local Similarity 57.1% Pred. No. 98;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 SDGLMNNQTOLE 15
DB 94 SDDMENTLSQFL 107

RESULT 21
5395759-2

Patent No. 5395759
APPLICANT: HOLMES, IAN H.; DYALL-SMITH, MICHAEL L.
TITLE OF INVENTION: DNA SEQUENCES AND AMINO ACID SEQUENCE
ENCODING THE HUMAN ROTAVIRUS MAJOR OUTER CAPSID GLYCOPROTEIN
NUMBER OF SEQUENCES: 14
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/474,642
FILING DATE: 29-APR-1985
PRIOR APPLICATION NUMBER: 824,704
APPLICATION NUMBER: 824,704
FILING DATE: 04-FEB-1987
SEQ ID NO: 2
LENGTH: 326
5395759-2

Query Match 41.4% Score 41; DB 6; Length 326;
Best Local Similarity 57.1% Pred. No. 98;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 SDGLMNNQTOLE 15
DB 94 SDDMENTLSQFL 107

RESULT 22
US-09-248-796A-18622
Sequence 18622, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstein et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 18622
LENGTH: 365
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-18622

Query Match 41.4% Score 41; DB 4; Length 365;
Best Local Similarity 53.8% Pred. No. 1.1e+02;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 GLMNNQTOLE 16
DB 279 GLMNNATDFNIE 291

RESULT 23
US-09-248-796A-17089
Sequence 17089, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstein et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 17089
LENGTH: 817

;; TYPE: PRT
;; ORGANISM: Candida albicans
US-09-248-796A-17089

Query Match 41.4%; Score 41; DB 4; Length 817;
Best Local Similarity 57.1%; Pred. NO. 2.7e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 SSDGLMNNOTOLF 14
|||:|:
Db 327 SSDGLMNDNPRYLF 340

RESULT 24
US-08-840-062-5
; Sequence 5, Application US/08840062
; Patent No. 6117977
; GENERAL INFORMATION:
; APPLICANT: LASKY, LAURENCE A.
; APPLICANT: WU, KAI
; TITLE OF INVENTION: TYPE C LECTINS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winpatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/840,062
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1019R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1455 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-840-062-5

Query Match 41.4%; Score 41; DB 3; Length 1455;
Best Local Similarity 87.5%; Pred. NO. 5.2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSDGLMNN 8
|||:
Db 1335 SSSGLMNN 1342

RESULT 25
US-08-231-193A-11
; Sequence 11, Application US/08231193A
; Patent No. 5649895
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorrie P.
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lu, Chin-Chun
; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR
; SUBUNITS, NUCLEIC ACIDS ENCODING SAME AND USES THEREFOR

;; NUMBER OF SEQUENCES: 63
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Brown, Martin, Haller & McClain
;; STREET: 1660 Union Street
;; CITY: San Diego
;; STATE: CA
;; COUNTRY: U.S.A.
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,193A
; FILING DATE: 20-APR-1994
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/052,459
; FILING DATE: 20-APR-1993
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-9383
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0999
; TELEFAX: 619-238-0062
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1464 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-231-193A-11

Query Match 41.4%; Score 41; DB 2; Length 1464;
Best Local Similarity 41.2%; Pred. NO. 5.3e+02;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSDGLMNNOTOLFLEH 17
:|:|:|:|:|:
Db 1174 NEGLSNDQKLYSKH 1190

RESULT 26
US-08-486-273A-11
; Sequence 11, Application US/08486273A
; Patent No. 5985586
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorrie P.
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lu, Chin-Chun
; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR SUBUNITS, DNA
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,273A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/231,193
FILING DATE: 20-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-9383B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-238-0999
TELEFAX: 619-238-0062
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1464 amino acids
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-486-273A-11

Query Match 41.4%; Score 41; DB 2; Length 1464;
Best Local Similarity 41.2%; Pred. No. 5.3e+02;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

1 SSDGLMNNQTOFLFH 17
:|||||:|:|:|
Db 1174 NEBGLSNNDOYKLYSKH 1190

RESULT 27
US-08-940-086A-11
Sequence 11, Application US/08940086A
Patent No. 6111091
GENERAL INFORMATION:
APPLICANT: Daggett, Lorrie P.
APPLICANT: Ellis, Steven B.
APPLICANT: Liaw, Chen W.
APPLICANT: Lu, Chih-Chun
TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR
TITLE OF INVENTION: SUBUNITS, NUCLEIC ACIDS ENCODING SAME AND USES THEREFOR
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/940,086A
FILING DATE: 29-SEPT-97
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/231,193
FILING DATE: 20-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/052,449
FILING DATE: 20-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24735-9383C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 450-8400
TELEFAX: (619) 450-8499
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1464 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-940-086A-11

Query Match 41.4%; Score 41; DB 3; Length 1464;
Best Local Similarity 41.2%; Pred. No. 5.3e+02;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

1 SSDGLMNNQTOFLFH 17
:|||||:|:|:|
Db 1174 NEBGLSNNDOYKLYSKH 1190

RESULT 28
US-08-436-332B-10
Sequence 10, Application US/08436332B
Patent No. 6130058
GENERAL INFORMATION:
APPLICANT: LE BOURDELLES, BEATRICE
APPLICANT: MYERS, BEATRICE
APPLICANT: WHITING, PAUL JOHN
TITLE OF INVENTION: STRABLY TRANSFECTED CELL LINE EXPRESSING
TITLE OF INVENTION: NMDA RECEPTORS, AND NOVEL CLONED NMDA
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: MICHAEL D. YABLONSKY - MERCK & CO., INC.
STREET: 126 EAST LINCOLN AVE., - P.O. BOX 2000
CITY: RAHWAY
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/436,332B
FILING DATE: 10-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: YABLONSKY, MICHAEL D.
REGISTRATION NUMBER: 40,407
REFERENCE/DOCKET NUMBER: T1210Y
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-4678
TELEFAX: 732-594-4720
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1464 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-436-332B-10

Query Match 41.4%; Score 41; DB 3; Length 1464;
Best Local Similarity 41.2%; Pred. No. 5.3e+02;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

1 SSDGLMNNQTOFLFH 17
:|||||:|:~|:|
Db 1174 NEBGLSNNDOYKLYSKH 1190

RESULT 29
US-08-940-035A-11
Sequence 11, Application US/08940035A
Patent No. 6316611
GENERAL INFORMATION:
APPLICANT: Daggett, Lorrie P.
APPLICANT: Ellis, Steven B.

APPLICANT: Liaw, Chen W.
APPLICANT: Lu, Chin-Chun
TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR
TITLE OF INVENTION: SUBUNITS, NUCLEIC ACIDS ENCODING SAME AND USES THEREFOR
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: U.S.A.
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/940,035A
FILING DATE: 29-SEPT-97
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/231,193
FILING DATE: 20-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/052,449
FILING DATE: 20-APR-1993
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-9383E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-238-0999
TELEFAX: 619-238-0062
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1464 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-940-035A-11

Query Match 41.4%; Score 41; DB 3; Length 1464;
Best Local Similarity 41.2%; Pred. No. 5.3e+02;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSDGLMNNNOTQFLFH 17
Db 1174 NEEGLSNNDQYKLYSKH 1190

RESULT 30
US-08-935-105A-11
Sequence 11, Application US/08935105A
Patent No. 6376660
GENERAL INFORMATION:
APPLICANT: Daggett, Lorrie P.
APPLICANT: Lu, Chin-Chun
TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR
TITLE OF INVENTION: SUBUNITS, NUCLEIC ACIDS ENCODING SAME AND USES THEREFOR
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: U.S.A.
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/935,105A
FILING DATE: 29-SEPT-97
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/231,193
FILING DATE: 20-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/052,449
FILING DATE: 20-APR-1993
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-9383D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-238-0999
TELEFAX: 619-238-0062
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1464 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-935-105A-11

Query Match 41.4%; Score 41; DB 3; Length 1464;
Best Local Similarity 41.2%; Pred. No. 5.3e+02;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSDGLMNNNOTQFLFH 17
Db 1174 NEEGLSNNDQYKLYSKH 1190

RESULT 31
US-09-648-797-11
Sequence 11, Application US/09648797
Patent No. 6469142
GENERAL INFORMATION:
APPLICANT: Daggett, Lorrie P.
APPLICANT: Liaw, Chen W.
APPLICANT: Lu, Chin-Chun
TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR
TITLE OF INVENTION: SUBUNITS, NUCLEIC ACIDS ENCODING SAME AND USES THEREFOR
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/648,797
FILING DATE: 28-AUG-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/940,086A
FILING DATE: 29-SEPT-97
APPLICATION NUMBER: US 08/231,193
FILING DATE: 20-APR-1994
APPLICATION NUMBER: US 08/052,449
FILING DATE: 20-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24735-9383C
TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 450-8400
TELEFAX: (619) 450-8499
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1464 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-648-797-11

Query Match 41.4%; Score 41; DB 4; Length 1464;
Best Local Similarity 41.2%; Pred. No. 5.3e+02;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 1 SSDGLMNNQOTQFLFLEH 17
: ||| ||| : : : :
Db 1174 NEEGLSNNDQYKLYSKH 1190

RESULT 32
US-08-217-704C-2
Sequence 2, Application US/08217704C
Patent No. 6489124
GENERAL INFORMATION:
APPLICANT: Folds, Robert
Fantaske, Robert
Adams, Sally-Lyn
Kamboj, Rajender
TITLE OF INVENTION: MODULATORY PROTEINS OF HUMAN CNS
RECEPTORS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: Suite 500, 3000 K Street
CITY: Washington, D.C., N.W.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/217,704C
FILING DATE: 25-Mar-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/987,953
FILING DATE: 11-DEC-94
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16777/259/ALIE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1464 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-217-704C-2

Query Match 41.4%; Score 41; DB 4; Length 1464;
Best Local Similarity 41.2%; Pred. No. 5.3e+02;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
Qy 1 SSDGLMNNQOTQFLFLEH 17
: ||| ||| : : : :
Db 1174 NEEGLSNNDQYKLYSKH 1190

RESULT 33 ..b.
US-09-386-123-11
Sequence 11, Application US/09386123
Patent No. 6521413
GENERAL INFORMATION:
APPLICANT: Daggett, Lorrie P.
APPLICANT: Ellis, Steven B.
APPLICANT: Liaw, Chen W.
APPLICANT: Lu, Chin-Chun
TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR
TITLE OF INVENTION: SUBUNITS, NUCLEIC ACIDS ENCODING SAME AND USES THEREFOR
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McNuliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: U.S.A.
ZIP: 92037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/386,123
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/486,273
FILING DATE: 06-JUNE-95
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/231,193
FILING DATE: 20-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/052,449
FILING DATE: 20-APR-1993
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-9383F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 858-450-8403
TELEFAX: 858-587-5360
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1464 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-386-123-11

Query Match 41.4%; Score 41; DB 4; Length 1464;
Best Local Similarity 41.2%; Pred. No. 5.3e+02;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
Qy 1 SSDGLMNNQOTQFLFLEH 17
: ||| ||| : : : :
Db 1174 NEEGLSNNDQYKLYSKH 1190

RESULT 34
US-09-187-049-8
Sequence 8, Application US/09187049
Patent No. 6117666
GENERAL INFORMATION:
APPLICANT: Lampka, Gayle K.
TITLE OF INVENTION: PLASTID PROTEOLYTIC PROCESSING ENZYME
TITLE OF INVENTION: THAT CLEAVES PRECURSOR POLYPEPTIDES
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:

```
/ ADDRESS: BRINKS HOFER GILSON & LIONE
/ STREET: P.O. Box 10395
/ CITY: Chicago
/ STATE: IL
/ COUNTRY: USA
/ ZIP: 60610
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/187,049
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/695,177
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Martin, Alice O.
/ REGISTRATION NUMBER: 35,601
/ REFERENCE/DOCKET NUMBER: 7814/16
/ TELEPHONE: 312 321-4200
/ TELEFAX: 312 321-4299
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 124 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ ORIGINAL SOURCE:
/ ORGANISM: MPP B Subunit from rat
/ US-09-187-049-8

Query Match 40.4%; Score 40; DB 3; Length 124;
Best Local Similarity 41.7%; Pred. No. 48;
Matches 10; Conservative 0; Mismatches 4; Indels 10; Gaps 1;

Cy 4 GLW-----NNNOTQLEH 17
Db 20 GLWIDAGSRYENKXNGTAHFLH 43

RESULT 35
US-09-540-236-3527
/ Sequence 3527, Application US/09540236
/ Patent No. 6673910
/ GENERAL INFORMATION:
/ APPLICANT: Gary L. Breton et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
/ FILE REFERENCE: 2709.2005-001
/ CURRENT APPLICATION NUMBER: US/09/540,236
/ CURRENT FILING DATE: 2000-04-04
/ NUMBER OF SEQ ID NOS: 3840
/ SEQ ID NO 3527
/ LENGTH: 126
/ TYPE: PRT
/ ORGANISM: M.catarhalis
/ US-09-540-236-3527

Query Match 40.4%; Score 40; DB 4; Length 126;
Best Local Similarity 43.8%; Pred. No. 48;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
```

```
Cy 3 DGLMNNQTOLEH 18
Db 95 DGVYTTNTGQDFVH 110

RESULT 36
```

```
US-09-513-999C-5695
/ Sequence 5695, Application US/09513999C
/ Patent No. 6783961
/ GENERAL INFORMATION:
/ APPLICANT: Dumas Milne Edwards, J.B.
/ APPLICANT: Duclert, A. Y.
/ APPLICANT: Giordano, J. Y.
/ TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
/ FILE REFERENCE: 59, US2, REG
/ CURRENT APPLICATION NUMBER: US/09/513,999C
/ CURRENT FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/122,487
/ PRIOR FILING DATE: 1999-02-26
/ NUMBER OF SEQ ID NOS: 36681
/ SOFTWARE: Patent.pm
/ SEQ ID NO 5695
/ LENGTH: 141
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-09-513-999C-5695

Query Match 40.4%; Score 40; DB 4; Length 141;
Best Local Similarity 41.7%; Pred. No. 55;
Matches 10; Conservative 0; Mismatches 4; Indels 10; Gaps 1;

Cy 4 GLW-----NNNOTQLEH 17
Db 82 GLWIDAGSRYENKXNGTAHFLH 105

RESULT 37
US-09-248-796A-16186
/ Sequence 16186, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Kelch Weinstock et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ CURRENT FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 16186
/ LENGTH: 186
/ TYPE: PRT
/ ORGANISM: Candida albicans
/ US-09-248-796A-16186

Query Match 40.4%; Score 40; DB 4; Length 186;
Best Local Similarity 66.7%; Pred. No. 75;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Cy 3 DGLMNNQTOLE 14
Db 34 DGLTNSYDQDLF 45

RESULT 38
US-08-807-263-4
/ Sequence 4, Application US/08807263C
/ Patent No. 5985627
/ GENERAL INFORMATION:
/ APPLICANT: Mortensen, Uffe
/ APPLICANT: Olesen, Kjeld
/ APPLICANT: Steniche, Henning
/ APPLICANT: Sorensen, Steen B.
/ APPLICANT: Bredam, Klaus
/ TITLE OF INVENTION: MODIFIED CARBOXYPEPTIDASE
```

```
FILE REFERENCE: 8648.71u801-no4
CURRENT APPLICATION NUMBER: US/08/807.263C
CURRENT FILING DATE: 1997-02-28
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 421
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-08-807-263-4

Query March 40.4%; Score 40; DB 2; Length 421;
Best Local Similarity 54.5%; Pred. No. 1.9e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 WNNNOTQLE 16
DB 84 WNSNATVIFLD 94

RESULT 39
US-09-375-140-10
Sequence 10, Application US/09375140
Patent No. 6489540
GENERAL INFORMATION:
APPLICANT: Kavanagh, T.
TITLE OF INVENTION: A NOVEL PLASTID-TARGETING NUCLEIC ACID SEQUENCE, A
TITLE OF INVENTION: NOVEL BETA-AMYLASE SEQUENCE, A STIMULUS-RESPONSIVE
FILE REFERENCE: 9341-017
CURRENT APPLICATION NUMBER: US/09/375.140
CURRENT FILING DATE: 1999-08-16
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10
LENGTH: 463
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-375-140-10

Query Match 40.4%; Score 40; DB 4; Length 463;
Best Local Similarity 42.9%; Pred. No. 2.1e+02;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 DGLMNNNOTQLE 16
DB 242 DGTWNSXGKPFME 255

RESULT 40
US-08-895-521-1
Sequence 1, Application US/08895521
Patent No. 5869311
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Shah, Purvi
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: MITOCHONDRIAL PROCESSING PEPTIDASE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
```

```
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/895.521
FILING DATE: Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0343 US
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 489 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT14
CLONE: 1593490
US-08-895-521-1

Query Match 40.4%; Score 40; DB 2; Length 489;
Best Local Similarity 41.7%; Pred. No. 2.2e+02;
Matches 10; Conservative 0; Mismatches 4; Indels 10; Gaps 1;

QY 4 GLM-----NNNOTQLEH 17
DB 82 GLMIDAGRYENKNGTAHLEH 105
```

Search completed: December 30, 2004, 16:14:38
Job time : 40 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 16:03:35 ; Search time 190 Seconds
(without alignments)
54.509 Million cell updates/sec

Title: US-10-718-321-1
Perfect score: 99
Sequence: 1 SSDLNNNNQTLFLHS 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : Uniprot 02:.*
1: uniprot_sprot:.*
2: uniprot_trembl:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99	100.0	359	2	043656 homo sapien
2	99	100.0	364	2	096d42 homo sapien
3	99	100.0	364	2	096d42 homo sapien
4	84	84.8	451	2	095144 cercoptithec
5	84	84.8	460	2	018984 cercoptithec
6	84	84.8	469	2	073748 cercoptithec
7	84	84.8	473	2	073747 cercoptithec
8	84	84.8	474	2	046597 cercoptithec
9	84	84.8	478	2	046598 cercoptithec
10	51	51.5	392	1	CEMA_NEPOL
11	49	49.5	267	1	082598 nephrobelm
12	49	49.5	278	1	CEMA_GUTH
13	49	49.5	338	2	P73105 synecocyst
14	48	48.5	129	2	068RM1 mus musculu
15	48	48.5	416	1	HAPC_PIG
16	48	48.5	558	2	036293 venezuelan
17	48	48.5	558	2	036294 venezuelan
18	48	48.5	559	2	036297 venezuelan
19	48	48.5	559	2	036298 venezuelan
20	48	48.5	559	2	036308 venezuelan
21	48	48.5	559	2	036309 venezuelan
22	48	48.5	559	2	041997 venezuelan
23	48	48.5	559	2	09WCF9 venezuelan
24	48	48.5	559	2	09WCG0 venezuelan
25	48	48.5	559	2	098755 venezuelan
26	48	48.5	559	2	098757 venezuelan
27	48	48.5	559	2	098758 venezuelan
28	48	48.5	1254	1	POL5_BEVY8
29	48	48.5	1254	1	POL5_BEVY8
30	48	48.5	1254	1	POL5_BEVY8
31	48	48.5	1254	2	Q77V76 venezuelan

32	48	48.5	1254	2	Q9YK44 venezuelan
33	48	48.5	1254	2	AAD14551 venezuela
34	48	48.5	1255	1	POL5_BEVY3
35	48	48.5	1255	1	POL5_BEVY3
36	48	48.5	1255	1	POL5_BEVY3
37	48	48.5	1255	2	036287 venezuelan
38	48	48.5	1255	2	036288 venezuelan
39	48	48.5	1255	2	036300 venezuelan
40	48	48.5	1255	2	090164 venezuelan
41	48	48.5	1255	2	091628 venezuelan
42	48	48.5	1255	2	066593 venezuelan
43	48	48.5	1255	2	066595 venezuelan
44	48	48.5	1255	2	09WCF4 venezuelan
45	48	48.5	1255	2	09WCF6 venezuelan
46	48	48.5	1255	2	09WCF7 venezuelan
47	48	48.5	1255	2	09WCF8 venezuelan
48	48	48.5	1255	2	09WCF9 venezuelan
49	48	48.5	1255	2	098773 venezuelan
50	48	48.5	1255	2	098775 venezuelan
51	48	48.5	1259	2	Q9YK55 venezuelan
52	47	47.5	67	2	Q7YD94 echinops te
53	47	47.5	68	2	Q9G389 echinops te
54	47	47.5	190	2	Q6CB54 yartowia li
55	47	47.5	1258	2	Q9WC27 venezuelan
56	47	47.5	1404	2	Q9T011 arabidopsis
57	46	46.5	446	2	Q6CUB0 kluyveromyc
58	46	46.5	203	2	Q75020 human immun
59	46	46.5	203	2	Q75021 human immun
60	46	46.5	204	2	Q75022 human immun
61	46	46.5	326	2	Q9WN97 human rolav
62	46	46.5	508	1	YBY9_YEAST
63	46	46.5	545	2	Q94EJ9 saccharomyc
64	45	45.0	809	2	Q8S159 oryza sativ
65	45	45.5	230	2	Q9C6X3 arabidopsis
66	45	45.5	272	2	Q9H7M5 homo sapien
67	45	45.5	273	2	Q821N7 streptomyce
68	45	45.5	301	1	LEC1 RAT
69	45	45.5	309	2	Q9NMG8 rattus norv
70	45	45.5	470	2	Q9NMG8 rattus norv
71	45	45.5	470	2	Q9NMG8 rattus norv
72	45	45.5	687	2	Q9XDI8 rattus norv
73	45	45.5	729	2	Q88K44 pseudomonas
74	45	45.5	1464	1	Q6FLH3 candida gla
75	45	45.5	1464	1	Q6FLH3 candida gla
76	45	45.5	1464	1	Q6FLH3 candida gla
77	45	45.5	1464	1	Q6FLH3 candida gla
78	45	45.5	1464	1	Q6FLH3 candida gla
79	44	44.9	329	2	Q92M89 rattus norv
80	44	44.4	155	2	Q9XSL8 ovis arie
81	44	44.4	202	2	Q75012 human immun
82	44	44.4	205	2	Q75014 human immun
83	44	44.4	207	2	Q75015 human immun
84	44	44.4	236	1	CEMA_MESV1
85	44	44.4	250	2	Q88615 meriones un
86	44	44.4	253	2	Q72N84 leptospira
87	44	44.4	253	2	Q88FC3 leptospira
88	44	44.4	253	2	Q88FC3 leptospira
89	44	44.4	272	2	Q7VOK4 candidatus
90	44	44.4	359	2	Q6G3Z8 bartonella
91	44	44.4	391	2	Q6G164 bartonella
92	44	44.4	429	2	Q82VJ3 enterococcu
93	44	44.4	449	1	CEMA_PSINU
94	44	44.4	474	2	Q72FV1 desulfovibr
95	44	44.4	474	2	Q72FV1 desulfovibr
96	44	44.4	488	2	Q6CDV9 yartowia li
97	44	44.4	524	2	Q757J1 ashbya gos
98	44	44.4	524	2	Q757J1 ashbya gos
99	44	44.4	554	2	Q757J1 ashbya gos
100	44	44.4	564	2	Q757J1 ashbya gos
101	44	44.4	588	2	Q6L984 homo sapien
102	44	44.4	588	2	Q6L984 homo sapien
103	44	44.4	589	1	Q9ERT3 mus musculu
104	44	44.4	589	1	Q9ERT3 mus musculu

```

105 44 44.4 589 2 BAC38088 BAC38088 mus muscu
106 44 44.4 590 1 ACM3_BOVIN P41984 bos taurus
107 44 44.4 590 1 ACM3_GORGO Q912a3 gorilla gor
108 44 44.4 590 1 ACM3_HUMAN P20309 homo sapien
109 44 44.4 590 1 ACM3_PANTR Q912a4 pan troglod
110 44 44.4 590 1 ACM3_PIG P11483 sus scrofa
111 44 44.4 590 1 ACM3_PONNY Q912a2 pongo pygma
112 44 44.4 639 1 ACM3_CHICK P49578 gallus galli
113 44 44.4 664 2 O68HW2 O68HW2 dictyosteli
114 44 44.4 681 2 O9A0S5 O9A0S5 bacillus sp
115 44 44.4 964 2 O47170 O47170 escherichia
116 44 44.4 973 2 O66728 O66728 streptomyc
117 44 44.4 1028 2 Q7BZV0 Q7BZV0 shigella fl
118 44 44.4 1028 1 BGA2_ECOLI P06864 escherichia
119 44 44.4 1030 1 BGA2_ECOLI P06864 escherichia
120 44 44.4 1030 2 O6B5F0 O6B5F0 escherichia
121 44 44.4 1042 2 O7AP0 O7AP0 escherichia
122 44 44.4 1042 2 O8XAM9 O8XAM9 escherichia
123 44 44.4 1084 2 O8PFI8 O8PFI8 ureaplasma
124 44 44.4 1089 2 O8PFI0 O8PFI0 escherichia
125 44 44.4 1338 2 O8A3N1 O8A3N1 bacteroides
126 44 44.4 2189 2 O8IKV6 O8IKV6 plasmodium
127 43 43.4 115 2 O90760 O90760 human immun
128 43 43.4 118 2 O8HDB7 O8HDB7 carteria ob
129 43 43.4 171 2 Q70Q29 Q70Q29 human immun
130 43 43.4 171 2 CAd87143 CAd87143 human imm
131 43 43.4 227 2 O9GH46 O9GH46 basichlamys
132 43 43.4 326 2 P74730 P74730 synechocyst
133 43 43.4 345 2 O67488 O67488 rotavirus a
134 43 43.4 345 2 O9T939 O9T939 shinisaurus
135 43 43.4 464 2 Q7XPC8 Q7XPC8 oryza sativ
136 43 43.4 464 2 Q76H39 Q76H39 chlamydomon
137 43 43.4 464 2 O8HDB6 O8HDB6 pseudocarte
138 43 43.4 464 2 O8HDB9 O8HDB9 carteria ra
139 43 43.4 464 2 O8HDD1 O8HDD1 chlamydomon
140 43 43.4 464 2 O8HDD2 O8HDD2 chlamydomon
141 43 43.4 464 2 O8HDD6 O8HDD6 pediatrum
142 43 43.4 464 2 BAC87678 BAC87678 chlamydom
143 43 43.4 484 2 O9U9P0 O9U9P0 hydra atten
144 43 43.4 498 2 O8LIU6 O8LIU6 vitreochlam
145 43 43.4 498 2 O8MPF7 O8MPF7 vitreochlam
146 43 43.4 498 2 O8GH48 O8GH48 tetraena
147 43 43.4 516 2 Q7SY19 Q7SY19 brachydano
148 43 43.4 525 2 Q7PYN5 Q7PYN5 anopheles g
149 43 43.4 579 2 O6PDM4 O6PDM4 mus muscu
150 43 43.4 579 2 AAH58624 AAH58624 mus muscu

```

ALIGNMENTS

```

RESULT 1
ID 043656 PRELIMINARY; PRT; 359 AA.
AC 043656;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-JUN-2004 (TREMBlrel. 26, Last annotation update)
DE Hepatitis A virus cellular receptor 1.
GN Name=HAVcr-1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=98325180; PubMed=9658108;
RA Feigelshtock D., Thompson P., Mattoo P., Zhang Y., Kaplan G.G.;
RT "The human homolog of HAVcr-1 codes for a hepatitis A virus cellular
RT receptor.";
RL J. Virol. 72:6621-6628(1998).
EMBL AF043724; AAC39862.1; -.

```

```

DR GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00409; Ig; 1.
DR PROSITE; PS50835; Ig_LIKE; 1.
DR PROSITE; PS00290; Ig_MHC; UNKNOWN_1.
DR RECEPTOR.
SO SEQUENCE 359 AA; 38704 MW; C207PEC562DC62CA CRC64;

Query Match 100.0%; Score 99; DB 2; Length 359;
Best Local Similarity 100.0%; Pred. No. 3.1e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SSDGLNNNOTOLFLEHS 18
Db 264 SSDGLNNNOTOLFLEHS 281

RESULT 2
ID 096D42 PRELIMINARY; PRT; 364 AA.
AC 096D42;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE HAVCR1 protein.
GN Name=HAVCR1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22386257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heintz F.,
RA Ditschenko L., Marziska K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhally S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Nadeau A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Boultard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,
RA Krzywinski M.I., Skaleck U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RN Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RN [3]
RP Ebert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC013325; AAH13325.1; -.
DR EMBL; CR457114; CAG33395.1; -.
DR Genew; HGNC:17866; HAVCR1.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.

```

DR SMART: SM00409; IG, 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 364 AA; 39249 MW; AA67C7DC7FAC81F1 CRC64;

Query Match 100.0%; Score 99; DB 2; Length 364;
Best Local Similarity 100.0%; Pred. No. 3.2e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSDGLMNNQOTQFLFHS 18
DB 269 SSDGLMNNQOTQFLFHS 286

RESULT 3
CAG33395 PRELIMINARY; PRT; 364 AA.
ID CAG33395;
AC CAG33395;

DT 01-JUN-2004 (TREMBLrel. 27, Created)
DT 01-JUN-2004 (TREMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TREMBLrel. 27, Last annotation update)
DE HAVCR1 protein.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;

RN [1]
RP SEQUENCE FROM N.A.
RA Ebert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;
RT "Cloning of human full open reading frames in Gateway(TM) system entry
RT vector (pDONR201)"; to the EMBL/GenBank/DBJ databases.
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR457114; CAG33395.1; -.

SQ SEQUENCE 364 AA; 39249 MW; AA67C7DC7FAC81F1 CRC64;

Query Match 100.0%; Score 99; DB 2; Length 364;
Best Local Similarity 100.0%; Pred. No. 3.2e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSDGLMNNQOTQFLFHS 18
DB 269 SSDGLMNNQOTQFLFHS 286

RESULT 4

O95144 PRELIMINARY; PRT; 451 AA.
ID O95144;
AC O95144;

DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE HAVCR-1 protein precursor.

OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;

RN [1]
RP SEQUENCE FROM N.A.
RA Tissue:Kidney;
RX MEDLINE=97015129; PubMed=8861957;
RA Kaplan G., Totsuka A., Thompson P., Akatsuka T., Moritsugu Y.,
RA Feinstein S.M.;

RT "Identification of a surface glycoprotein on African green monkey
RT kidney cells as a receptor for hepatitis A virus";
RL EMO J. 15:4282-4296(1996).

DR EMBL; X98952; CAA66906.1; -.
DR PIR; S71754; S71754.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG_LIKE.
DR InterPro; IPR003006; IG_MHC.
DR SMART; SM00409; IG, 1.

DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Signal.
FT SIGNAL 17 Potential.
SQ SEQUENCE 451 AA; 48774 MW; 5D395C5455AA4332 CRC64;

Query Match 84.8%; Score 84; DB 2; Length 451;
Best Local Similarity 88.9%; Pred. No. 9.5e-05;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SSDGLMNNQOTQFLFHS 18
DB 344 SSDGLMNNQOTQFLFHS 361

RESULT 5
O18984 PRELIMINARY; PRT; 460 AA.
ID O18984;
AC O18984;

DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Hepatitis A virus receptor.

OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97368416; PubMed=9225030;
RA Ashida M., Hamada C.;

RT "Molecular cloning of the hepatitis A virus receptor from a simian
RT cell line";
RL J. Gen. Virol. 78:1565-1569(1997).
DR EMBL; D88585; BAA21556.1; -.

DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG_LIKE.
DR InterPro; IPR003006; IG_MHC.
DR SMART; SM00409; IG, 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.

SQ SEQUENCE 460 AA; 49687 MW; 1D1A0C1D832139EA CRC64;

Query Match 84.8%; Score 84; DB 2; Length 460;
Best Local Similarity 88.9%; Pred. No. 9.7e-05;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SSDGLMNNQOTQFLFHS 18
DB 353 SSDGLMNNQOTQFLFHS 370

RESULT 6

O70048 PRELIMINARY; PRT; 469 AA.
ID O70048;
AC O70048;

DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Hepatitis A virus cellular receptor 1 short form.

GN Name=HAVCR-1;
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98285791; PubMed=9621093;
RA Feigelshtock D., Thompson P., Mattoo P., Kaplan G.G.;

RT "Polymorphisms of the hepatitis A virus cellular receptor 1 in African
RT green monkey kidney cells result in antigenic variants that do not
RT react with protective monoclonal antibody 190/4.";
RL J. Virol. 72:6218-6222(1998).
DR EMBL: AF043446; AAC39771.1; -
DR GO: GO:0004872; F:receptor activity; IEA.
DR InterPro: IPR003599; IG.
DR InterPro: IPR007110; IG_1like.
DR SMART: SMART: IPR003006; IG_MHC.
DR SMART: SMART: IPR003006; IG_MHC.
DR PROSITE: PS50835; IG LIKE; 1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
KW Receptor.
SQ SEQUENCE 469 AA; 50561 MW; 27E386F47948F528 CRC64;
Query Match 84.8%; Score 84; DB 2; Length 469;
Best Local Similarity 88.9%; Pred. No. 9.9e-05; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 2;
QY 1 SSDGLMNNQOTLFLEHS 18
Db 362 SSDGLMNNQOTLFLEHS 379
RESULT 7
QY 07J347 PRELIMINARY; PRT; 473 AA.
AC 07J347;
DT 05-JUN-2004 (TREMBlrel. 27, Created)
DT 05-JUN-2004 (TREMBlrel. 27, Last sequence update)
DE Hepatitis A virus cellular receptor 1 short form.
GN Name=HAVcr-1;
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98285791; PubMed=9621093;
RA Feigelsstock D., Thompson P., Mattoo P., Kaplan G.G.;
RT "Polymorphisms of the hepatitis A virus cellular receptor 1 in African
RT green monkey kidney cells result in antigenic variants that do not
RT react with protective monoclonal antibody 190/4.";
RL J. Virol. 72:6218-6222(1998).
DR EMBL: AF043448; AAC39773.1; -
DR GO: GO:0004872; F:receptor activity; IEA.
DR InterPro: IPR003599; IG.
DR InterPro: IPR007110; IG_1like.
DR InterPro: IPR003006; IG_MHC.
DR SMART: SMART: IPR003006; IG_MHC.
DR PROSITE: PS50835; IG LIKE; 1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
KW Receptor.
SQ SEQUENCE 473 AA; 50973 MW; CD15F5EE79C8013 CRC64;
Query Match 84.8%; Score 84; DB 2; Length 473;
Best Local Similarity 88.9%; Pred. No. 0.0001;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 SSDGLMNNQOTLFLEHS 18
Db 366 SSDGLMNNQOTLFLEHS 383
RESULT 8
QY 046597 PRELIMINARY; PRT; 474 AA.
AC 046597;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)

DE Hepatitis A virus cellular receptor 1 long form.
GN Name=HAVcr-1;
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98285791; PubMed=9621093;
RA Feigelsstock D., Thompson P., Mattoo P., Kaplan G.G.;
RT "Polymorphisms of the hepatitis A virus cellular receptor 1 in African
RT green monkey kidney cells result in antigenic variants that do not
RT react with protective monoclonal antibody 190/4.";
RL J. Virol. 72:6218-6222(1998).
DR EMBL: AF043447; AAC39772.1; -
DR GO: GO:0004872; F:receptor activity; IEA.
DR InterPro: IPR003599; IG.
DR InterPro: IPR007110; IG_1like.
DR InterPro: IPR003006; IG_MHC.
DR SMART: SMART: IPR003006; IG_MHC.
DR PROSITE: PS50835; IG LIKE; 1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
KW Receptor.
SQ SEQUENCE 474 AA; 51088 MW; 472DF987CA220524 CRC64;
Query Match 84.8%; Score 84; DB 2; Length 474;
Best Local Similarity 88.9%; Pred. No. 0.0001;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 SSDGLMNNQOTLFLEHS 18
Db 367 SSDGLMNNQOTLFLEHS 384
RESULT 9
QY 046598 PRELIMINARY; PRT; 478 AA.
AC 046598;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Hepatitis A virus cellular receptor 1 long form.
GN Name=HAVcr-1;
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98285791; PubMed=9621093;
RA Feigelsstock D., Thompson P., Mattoo P., Kaplan G.G.;
RT "Polymorphisms of the hepatitis A virus cellular receptor 1 in African
RT green monkey kidney cells result in antigenic variants that do not
RT react with protective monoclonal antibody 190/4.";
RL J. Virol. 72:6218-6222(1998).
DR EMBL: AF043449; AAC39774.1; -
DR GO: GO:0004872; F:receptor activity; IEA.
DR InterPro: IPR003599; IG.
DR InterPro: IPR007110; IG_1like.
DR InterPro: IPR003006; IG_MHC.
DR SMART: SMART: IPR003006; IG_MHC.
DR PROSITE: PS50835; IG LIKE; 1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
KW Receptor.
SQ SEQUENCE 478 AA; 51501 MW; 3A919655C752FF60 CRC64;
Query Match 84.8%; Score 84; DB 2; Length 478;
Best Local Similarity 88.9%; Pred. No. 0.0001;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 SSDGLMNNQOTLFLEHS 18

Db 371 SSDGLMNNQTOLEPHS 388

```

RESULT 10
CEMA_NEPOL STANDARD; PRT; 392 AA.
ID CEMA_NEPOL
AC 09TKR22;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Chloroplast envelope membrane protein.
GN Name=Cema;
OS Nephroselmis olivacea.
OC Chloroplast.
OC Eukaryota, Viridiplantae; Chlorophyta; Prasinophyceae;
OC Chlorodendreales; Chlorodendraceae; Nephroselmis.
OX NCBI_TaxID=31312;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIES-484;
RA MEDLINE=99398694; PubMed=10468594;
RT Tunnel M.; Odis C.; Lemieux C.;
RT "The complete chloroplast DNA sequence of the green alga Nephroselmis
RT olivacea: insights into the architecture of ancestral chloroplast
RT genomes".
RT Proc. Natl. Acad. Sci. U.S.A. 96:10248-10253(1999).
RL -1- FUNCTION: May be involved in proton extrusion. Indirectly promotes
CC efficient inorganic carbon uptake into chloroplasts (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast inner
CC envelope (By similarity).
CC -1- SIMILARITY: Belongs to the cema family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF137379; AAC54824.1; -
DR HAMAP; MF_01308; -; 1.
DR InterPro; IPR004282; Cema.
DR Pfam; PF03040; Cema; 1.
KW Chloroplast; Hydrogen ion transport; Transmembrane; Transport.
FT TRANSMEM 172 194 Potential.
FT TRANSMEM 267 289 Potential.
FT TRANSMEM 315 337 Potential.
FT TRANSMEM 352 373 Potential.
SQ SEQUENCE 392 AA; 45537 MW; DVC9E74A56FEB888 CRC64;

Query Match 51.5%; Score 51; DB 1; Length 392;
Best Local Similarity 69.2%; Pred. No. 14;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 6 WNNQTOLEPHS 18
Db 202 WNNQTOLEPHS 214

RESULT 11
Q8ZS98 PRELIMINARY; PRT; 267 AA.
ID Q8ZS98;
AC Q8ZS98;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE A17613 protein.
GN OrderedLocustName=a17613;
OS Anabaena sp. (strain PCC 7120).
SQ Plasmid PCC7120beta.

```

```

OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T.; Nakamura Y.; Wolk C.P.; Kuritz T.; Sasamoto S.;
RA Matsubara A.; Iriyuchi M.; Ishikawa A.; Kawashima K.; Kimura T.;
RA Kishida Y.; Kohara M.; Matsumoto M.; Matsuno A.; Muraki A.;
RA Nakazaki N.; Shimo S.; Sugimoto M.; Takazawa M.; Yanada M.;
RA Yasuda M.; Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120."
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003602; BAB77256.1; -
DR PIR; AB2541; AB2541.
DR InterPro; IPR001395; Aldo/ket red.
DR PROSITE; PS00063; ALDO-KETO_REDUCTASE_3; UNKNOWN_1.
KW Complete proteome; Plasmid.
SQ SEQUENCE 267 AA; 31095 MW; 0CF635901AE8F5FE CRC64;

OY 2 SDGLMNNQTOLEPH 15
Db 190 SDGLMNNQTOLEPH 203

Query Match 49.5%; Score 49; DB 2; Length 267;
Best Local Similarity 57.1%; Pred. No. 18;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 2 SDGLMNNQTOLEPH 15
Db 190 SDGLMNNQTOLEPH 203

RESULT 12
CEMA_GUITH STANDARD; PRT; 278 AA.
ID CEMA_GUITH
AC 078470;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Chloroplast envelope membrane protein.
GN Name=cema; Synonyms=ycf10;
OS Guillardia theta (Cryptomonas phi).
OC Chloroplast.
OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
OX NCBI_TaxID=55529;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99128221; PubMed=9929392;
RA Douglas S.E.; Penny S.U.;
RT "The plastid genome of the cryptophyte alga, Guillardia theta:
RT complete sequence and conserved synteny groups confirm its common
RT ancestry with red algae."
RL J. Mol. Evol. 48:236-244(1999).
RL -1- FUNCTION: May be involved in proton extrusion. Indirectly promotes
CC efficient inorganic carbon uptake into chloroplasts (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast inner
CC envelope (By similarity).
CC -1- SIMILARITY: Belongs to the cema family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF041468; AAC35661.1; -
DR HAMAP; MF_01308; -; 1.
DR InterPro; IPR004282; Cema.
DR Pfam; PF03040; Cema; 1.
KW Chloroplast; Hydrogen ion transport; Transmembrane; Transport.
FT TRANSMEM 60 80 Potential.
FT TRANSMEM 163 183 Potential.
FT TRANSMEM 201 221 Potential.

```

```

FT TRANSMEM 239 259 Potential.
SQ SEQUENCE 278 AA; 32209 MW; 1E88817DC25582DE CRC64;
Query Match 49.5%; Score 49; DB 1; Length 278;
Best Local Similarity 50.0%; Pred. No. 19;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 3 DGLMNNQTOLEFHS 18
Db 85 DYFNNQSDIFLNS 100

RESULT 13
P73105 PRELIMINARY; PRT; 338 AA.
ID P73105
AC P73105;
DT 01-FEB-1997 (TRENBLrel. 02. Created)
DT 01-FEB-1997 (TRENBLrel. 02. Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24. Last annotation update)
DE N-acetylmuramoyl-L-alanine amidease.
GN Name=amiA; OrderedLocustNames=slr1910;
OS Synchocystis sp. (Strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugita M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synchocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=PCC6803;
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugita M., Sasamoto S., Kilmura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naro K., Okumura S.,
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synchocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
DR EMBL: D90903; BAI17131.1; -.
DR PIR: S75217; S75217.
DR GO: GO:0008745; F:N-acetylmuramoyl-L-alanine amidease activity; IEA.
DR GO: GO:0009253; P:peptidoglycan catabolism; IEA.
DR InterPro: IPR002508; Amidase_3_hydro.
DR Pfam: PF01520; Amidase_3; 1.
DR SMART: SMO0646; Ami_3; 1.
DR Complete proteome.
KW SEQUENCE 338 AA; 37929 MW; 0013BA7DDE297DC5 CRC64;

Query Match 49.5%; Score 49; DB 2; Length 338;
Best Local Similarity 61.5%; Pred. No. 24;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 4 GLMNNQTOLEFHS 16
Db 95 GRMNNSTRIVLE 107

RESULT 14
Q8BRW1 PRELIMINARY; PRT; 129 AA.
ID Q8BRW1
AC Q8BRW1;
DT 01-MAR-2003 (TRENBLrel. 23. Created)
DT 01-MAR-2003 (TRENBLrel. 23. Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25. Last annotation update)
DE Mus musculus adult male aorta and vein cDNA, RIKEN full-length

```

```

DE enriched library, clone:A530092U01 product:hypothetical protein, full
DE insert sequence.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=1085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kiteunai T., Tashiro H., Itoh M.,
RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishie T., Harada A.,
RA Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kaishogi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Matsubiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsumura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
RA Hori F., Imocani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ono M., Ohato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai K., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK041221; BAC30867.1; -.
DR Hypothetical protein.
KW SEQUENCE 129 AA; 14724 MW; A9D2912B874A4865 CRC64;

Query Match 48.5%; Score 48; DB 2; Length 129;
Best Local Similarity 56.2%; Pred. No. 12;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```

Qy 2 SDGLMNNNOTQLEFH 17
Db 52 SCGLMSPQTHLFIWH 67

RESULT 15

HAPC_PIG STANDARD; PRT; 416 AA.
AC P17630;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Haptocorin precursor (R protein) (Cobalophilin) (Fragment).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Stomach.
RX MEDLINE=90235846; PubMed=2331993;
RA Hewitt J.E., Seetharam B., Leykam J.F., Alpers D.H.;
RT "Isolation and characterization of a cDNA encoding porcine gastric
haptocorin.";
RL Eur. J. Biochem. 189:125-130(1990).
CC -1- FUNCTION: May play a role in preventing the absorption of
cobalamin analogs produced by bacteria. Binds to cobalamin and to
cobalamin analogs such as cobinamide.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Haptocorins are a family of cobalamin-binding
glycoproteins found in blood, salivary and mucosal secretions.
CC -1- PTM: Contains about 30% carbohydrates.
CC -1- SIMILARITY: Belongs to the eukaryotic cobalamin transport proteins
family.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----

DR EMBL; X52566; CAA36800.1; -.
DR PIR; S09334; S09334.
DR InterPro; IPR002157; Cobalamin bind.
DR Pfam; PF01122; Cobalamin bind; 1.
DR PROSITE; PS00468; COBALAMIN BINDING; 1.
KW Cobalt transport; Direct protein sequencing; Glycoprotein; Signal;
KW Transport.
FT NON_TER 1
FT SIGNAL 24
FT CHAIN 25
FT DISULFID 155
FT CARBOHYD 89
FT CARBOHYD 160
FT CARBOHYD 165
FT CARBOHYD 178
FT CARBOHYD 311
FT CARBOHYD 327
FT CARBOHYD 344
FT CARBOHYD 359
SQ SEQUENCE 416 AA; 46963 MW; 3380CC6F4C8E4AE3 CRC64;

Query Match 48.5%; Score 48; DB 1; Length 416;
Best Local Similarity 44.4%; Pred. No. 44;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 SSGLMNNNOTQLEFH 18
Db 378 SVGLMNNNNERTYWEHS 395

RESULT 16
ID 036293 PRELIMINARY; PRT; 558 AA.
AC 036293;
DT 01-JAN-1998 (TRENBLREL. 05, Created)
DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
DE Polypeptide (Fragment).
OS Venezuelan equine encephalitis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus; VEEV complex.
OX NCBI_TaxID=11036;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V209A;
RX MEDLINE=97404681; PubMed=9261393;
RA Powers A.M., Oberste M.S., Brault A.C., Rico-Hesse R., Schumura S.M.,
RA Smith J.F., Kang W., Sweeney W.P., Weaver S.C.;
RT "Repeated emergence of epidemic/epizootic Venezuelan equine
encephalitis from a single genotype of enzootic subtype ID virus.";
RL J. Virol. 71:6937-6705(1997).
DR EMBL; AF004464; AAB81710.1; -.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0019031; C:Viral envelope; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR002548; Alpha_E1_glycop.
DR InterPro; IPR000936; Alpha_E2_glycop.
DR InterPro; IPR002533; Alpha_E3_glycop.
DR InterPro; IPR000336; Flav_glycoprote.
DR Pfam; PF01589; Alpha_E1_glycop; 1.
DR Pfam; PF00943; Alpha_E2_glycop; 1.
DR Pfam; PF01563; Alpha_E3_glycop; 1.
KW Polypeptide.
FT NON_TER 1
FT CHAIN 537
FT CHAIN 538
FT NON_TER 558
SQ SEQUENCE 558 AA; 61777 MW; 82FCDS3D51D15729 CRC64;

Query Match 48.5%; Score 48; DB 2; Length 558;
Best Local Similarity 50.0%; Pred. No. 60;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SSGLMNNNOTQLEFH 16
Db 487 SLDLMNNNOQMEFIQ 502

RESULT 17
ID 036294 PRELIMINARY; PRT; 558 AA.
AC 036294;
DT 01-JAN-1998 (TRENBLREL. 05, Created)
DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
DE Polypeptide (Fragment).
OS Venezuelan equine encephalitis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus; VEEV complex.
OX NCBI_TaxID=11036;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V209A;
RX MEDLINE=97404681; PubMed=9261393;
RA Powers A.M., Oberste M.S., Brault A.C., Rico-Hesse R., Schumura S.M.,
RA Smith J.F., Kang W., Sweeney W.P., Weaver S.C.;
RT "Repeated emergence of epidemic/epizootic Venezuelan equine
encephalitis from a single genotype of enzootic subtype ID virus.";
RL J. Virol. 71:6937-6705(1997).
DR EMBL; AF004465; AAB81711.1; -.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0019031; C:Viral envelope; IEA.

```
DR GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR002548; Alpha_E1_glycop.
DR InterPro: IPR000936; Alpha_E2_glycop.
DR InterPro: IPR002533; Alpha_E3_glycop.
DR InterPro: IPR000336; Flavi_glycoprote.
DR Pfam: PF01589; Alpha_E1_glycop; 1.
DR Pfam: PF00943; Alpha_E2_glycop; 1.
DR Pfam: PF01563; Alpha_E3_glycop; 1.
KM Polypeptid.
FT NON_TER 1
FT CHAIN <1 537 envelope glycoprotein E2.
FT CHAIN 538 >558 envelope glycoprotein E1.
FT NON_TER 558
SQ SEQUENCE 558 AA; 61626 MW; 95ED71BDDF81B58C CRC64;

Query Match 48.5%; Score 48; DB 2; Length 558;
Best Local Similarity 50.0%; Pred. No. 60;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SSDGLMNNQOTQFLF 16
Db 487 SLDLHMNNQOMFWIQ 502

RESULT 18
ID 036297 PRELIMINARY; PRT; 559 AA.
AC 036297;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polypeptid (Fragment).
OS Venezuelan equine encephalitis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus; VEEV complex.
OX NCBI_TaxID=11036;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=59001;
RX MEDLINE=97404681; PubMed=9261393;
RA Powers A.M., Oberste M.S., Brault A.C., Rico-Hesse R., Schmura S.M.,
RA Smith J.F., Kang W., Sweeney W.P., Weaver S.C.;
RT "Repeated emergence of epidemic/epizootic Venezuelan equine
RT encephalitis from a single genotype of enzootic subtype ID virus.";
RL J. Virol. 71:6697-6705(1997).
DR EMBL: AF004468; AAB81713.1; -.
DR GO:00019028; C:viral capsid; IEA.
DR GO:0019031; C:viral envelope; IEA.
DR GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR002548; Alpha_E1_glycop.
DR InterPro: IPR000936; Alpha_E2_glycop.
DR InterPro: IPR002533; Alpha_E3_glycop.
DR InterPro: IPR000336; Flavi_glycoprote.
DR Pfam: PF01589; Alpha_E1_glycop; 1.
DR Pfam: PF00943; Alpha_E2_glycop; 1.
DR Pfam: PF01563; Alpha_E3_glycop; 1.
KM Polypeptid.
FT NON_TER 1
FT CHAIN <1 537 envelope glycoprotein E2.
FT CHAIN 538 >559 envelope glycoprotein E1.
FT NON_TER 559
SQ SEQUENCE 559 AA; 61837 MW; 2C273D694EDFB64D CRC64;

Query Match 48.5%; Score 48; DB 2; Length 559;
Best Local Similarity 50.0%; Pred. No. 61;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SSDGLMNNQOTQFLF 16
Db 487 SLDLHMNNQOMFWIQ 502

RESULT 19
```

```
036298
ID 036298 PRELIMINARY; PRT; 559 AA.
AC 036298;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polypeptid (Fragment).
OS Venezuelan equine encephalitis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus; VEEV complex.
OX NCBI_TaxID=11036;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pe5-47ec;
RX MEDLINE=97404681; PubMed=9261393;
RA Powers A.M., Oberste M.S., Brault A.C., Rico-Hesse R., Schmura S.M.,
RA Smith J.F., Kang W., Sweeney W.P., Weaver S.C.;
RT "Repeated emergence of epidemic/epizootic Venezuelan equine
RT encephalitis from a single genotype of enzootic subtype ID virus.";
RL J. Virol. 71:6697-6705(1997).
DR EMBL: AF004469; AAB81714.1; -.
DR GO:00019028; C:viral capsid; IEA.
DR GO:0019031; C:viral envelope; IEA.
DR GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR002548; Alpha_E1_glycop.
DR InterPro: IPR000936; Alpha_E2_glycop.
DR InterPro: IPR002533; Alpha_E3_glycop.
DR InterPro: IPR000336; Flavi_glycoprote.
DR Pfam: PF01589; Alpha_E1_glycop; 1.
DR Pfam: PF00943; Alpha_E2_glycop; 1.
DR Pfam: PF01563; Alpha_E3_glycop; 1.
KM Polypeptid.
FT NON_TER 1
FT CHAIN <1 537 envelope glycoprotein E2.
FT CHAIN 538 >559 envelope glycoprotein E1.
FT NON_TER 559
SQ SEQUENCE 559 AA; 61749 MW; 93A484BFC92217F CRC64;

Query Match 48.5%; Score 48; DB 2; Length 559;
Best Local Similarity 57.9%; Pred. No. 61;
Matches 11; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

Qy 1 SSDGLMNNQOTQFLF 15
Db 487 SLDLHMNNQOMFWTQLLI 505

RESULT 20
ID 036308 PRELIMINARY; PRT; 559 AA.
AC 036308;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polypeptid (Fragment).
OS Venezuelan equine encephalitis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus; VEEV complex.
OX NCBI_TaxID=11036;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DE15191;
RX MEDLINE=97404681; PubMed=9261393;
RA Powers A.M., Oberste M.S., Brault A.C., Rico-Hesse R., Schmura S.M.,
RA Smith J.F., Kang W., Sweeney W.P., Weaver S.C.;
RT "Repeated emergence of epidemic/epizootic Venezuelan equine
RT encephalitis from a single genotype of enzootic subtype ID virus.";
RL J. Virol. 71:6697-6705(1997).
DR EMBL: AF0044652; AAB81715.1; -.
DR GO:00019028; C:viral capsid; IEA.
DR GO:0019031; C:viral envelope; IEA.
DR GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR002548; Alpha_E1_glycop.
```

```
DR InterPro: IPR00936; Alpha_E2_glycop.
DR InterPro: IPR002533; Alpha_E3_glycop.
DR InterPro: IPR000336; Flavi_glycoprote.
DR Pfam: PF01589; Alpha_E1_glycop; 1.
DR Pfam: PF00943; Alpha_E2_glycop; 1.
DR Pfam: PF01563; Alpha_E3_glycop; 1.
DR Polyprotein.
KW NON_TER
FT CHAIN <1 537 1
FT CHAIN 538 >559 envelope glycoprotein E2.
FT NON_TER 559 envelope glycoprotein E1.
SQ SEQUENCE 559 AA; 61784 MW; AF95E19333F72337 CRC64;

Query Match 48.5%; Score 48; DB 2; Length 559;
Best Local Similarity 50.0%; Pred. No. 61;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SSDGLMNNNQTOFLFLE 16
Db 487 SLDLMLNNNQOMFWIQ 502

RESULT 21
ID 036309 PRELIMINARY; PRT; 559 AA.
AC 036309;
DT 01-JAN-1998 (TREMblrel. 05. Created)
DT 01-JAN-1998 (TREMblrel. 05. Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25. Last annotation update)
DE Polyprotein (Fragment).
OS Venezuelan equine encephalitis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus; VEEV complex.
OX NCBI_TaxID=11036;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SH5;
RX MEDLINE=97404681; PubMed=9261393;
RA Powers A.M., Oberste M.S., Brault A.C., Rico-Hesse R., Schumura S.M.,
RA Smith J.F., Kang W., Sweeney W.P., Weaver S.C.;
RT "Repeated emergence of epidemic/epizootic Venezuelan equine
RT encephalitis from a single genotype of enzootic subtype ID virus.";
RL J. Virol. 71:6697-6705(1997).
DR EMBL: AF004853; AAB81716.1; -.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0019031; C:viral envelope; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR002548; Alpha_E1_glycop.
DR InterPro: IPR000936; Alpha_E2_glycop.
DR InterPro: IPR002533; Alpha_E3_glycop.
DR InterPro: IPR000336; Flavi_glycoprote.
DR Pfam: PF01589; Alpha_E1_glycop; 1.
DR Pfam: PF00943; Alpha_E2_glycop; 1.
DR Pfam: PF01563; Alpha_E3_glycop; 1.
KW Polyprotein.
FT CHAIN <1 537 1
FT CHAIN 538 >559 envelope glycoprotein E2.
FT NON_TER 559 envelope glycoprotein E1.
SQ SEQUENCE 559 AA; 61897 MW; 49FDE5F3BECBF720 CRC64;

Query Match 48.5%; Score 48; DB 2; Length 559;
Best Local Similarity 50.0%; Pred. No. 61;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SSDGLMNNNQTOFLFLE 16
Db 487 SLDLMLNNNQOMFWIQ 502

RESULT 22
ID 041997 PRELIMINARY; PRT; 559 AA.
AC 041997;
DT 01-JAN-1997;
DT 01-JAN-1997;
DT 01-OCT-2003 (TREMblrel. 25. Last annotation update)
DE Structural polyprotein (Fragment).
OS Venezuelan equine encephalitis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus; VEEV complex.
OX NCBI_TaxID=11036;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Panaquire;
RX MEDLINE=96332169; PubMed=8709783;
RA Weaver S.C., Salas R., Rico-Hesse R., Ludwig G.V., Oberste M.S.,
RA Boehell J., Tesh R.B.;
RT "Re-emergence of epidemic Venezuelan equine encephalomyelitis in South
RT America. VEE Study Group.";
RL Lancet 348:436-440(1996).
DR [2]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Panaquire;
RA Weaver S.C.;
RX Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: U55346; AAC52613.1; -.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0019031; C:viral envelope; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR002548; Alpha_E1_glycop.
DR InterPro: IPR000936; Alpha_E2_glycop.
DR InterPro: IPR002533; Alpha_E3_glycop.
DR InterPro: IPR000336; Flavi_glycoprote.
DR Pfam: PF01589; Alpha_E1_glycop; 1.
DR Pfam: PF00943; Alpha_E2_glycop; 1.
DR Pfam: PF01563; Alpha_E3_glycop; 1.
KW Polyprotein.
FT CHAIN <1 537 1
FT CHAIN 538 >559 protein E2.
FT NON_TER 559 protein E1.
SQ SEQUENCE 559 AA; 61754 MW; C48D0AE2677BE6A0 CRC64;

Query Match 48.5%; Score 48; DB 2; Length 559;
Best Local Similarity 50.0%; Pred. No. 61;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SSDGLMNNNQTOFLFLE 16
Db 487 SLDLMLNNNQOMFWIQ 502

RESULT 23
ID 09WC99 PRELIMINARY; PRT; 559 AA.
AC 09WC99;
DT 01-NOV-1999 (TREMblrel. 12. Created)
DT 01-NOV-1999 (TREMblrel. 12. Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25. Last annotation update)
DE Structural polyprotein (Fragment).
OS Venezuelan equine encephalitis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus; VEEV complex.
OX NCBI_TaxID=11036;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pura;
RX MEDLINE=99394886; PubMed=10466974;
RA Weaver S.C., Pfeiffer M., Mariotti K., Kang W., Kinney R.M.;
RT "Genetic evidence for the origins of Venezuelan equine encephalitis
RT virus subtype IAB outbreaks.";
RL Am. J. Trop. Med. Hyg. 60:441-448(1999).
DR EMBL: AF093104; AAD37003.1; -.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0019031; C:viral envelope; IEA.
```

```
AC 041997;
DT 01-JAN-1998 (TREMblrel. 05. Created)
DT 01-JAN-1998 (TREMblrel. 05. Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25. Last annotation update)
DE Structural polyprotein (Fragment).
OS Venezuelan equine encephalitis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus; VEEV complex.
OX NCBI_TaxID=11036;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Panaquire;
RX MEDLINE=96332169; PubMed=8709783;
RA Weaver S.C., Salas R., Rico-Hesse R., Ludwig G.V., Oberste M.S.,
RA Boehell J., Tesh R.B.;
RT "Re-emergence of epidemic Venezuelan equine encephalomyelitis in South
RT America. VEE Study Group.";
RL Lancet 348:436-440(1996).
DR [2]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Panaquire;
RA Weaver S.C.;
RX Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: U55346; AAC52613.1; -.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0019031; C:viral envelope; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR002548; Alpha_E1_glycop.
DR InterPro: IPR000936; Alpha_E2_glycop.
DR InterPro: IPR002533; Alpha_E3_glycop.
DR InterPro: IPR000336; Flavi_glycoprote.
DR Pfam: PF01589; Alpha_E1_glycop; 1.
DR Pfam: PF00943; Alpha_E2_glycop; 1.
DR Pfam: PF01563; Alpha_E3_glycop; 1.
KW Polyprotein.
FT CHAIN <1 537 1
FT CHAIN 538 >559 protein E2.
FT NON_TER 559 protein E1.
SQ SEQUENCE 559 AA; 61754 MW; C48D0AE2677BE6A0 CRC64;

Query Match 48.5%; Score 48; DB 2; Length 559;
Best Local Similarity 50.0%; Pred. No. 61;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SSDGLMNNNQTOFLFLE 16
Db 487 SLDLMLNNNQOMFWIQ 502

RESULT 23
ID 09WC99 PRELIMINARY; PRT; 559 AA.
AC 09WC99;
DT 01-NOV-1999 (TREMblrel. 12. Created)
DT 01-NOV-1999 (TREMblrel. 12. Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25. Last annotation update)
DE Structural polyprotein (Fragment).
OS Venezuelan equine encephalitis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus; VEEV complex.
OX NCBI_TaxID=11036;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pura;
RX MEDLINE=99394886; PubMed=10466974;
RA Weaver S.C., Pfeiffer M., Mariotti K., Kang W., Kinney R.M.;
RT "Genetic evidence for the origins of Venezuelan equine encephalitis
RT virus subtype IAB outbreaks.";
RL Am. J. Trop. Med. Hyg. 60:441-448(1999).
DR EMBL: AF093104; AAD37003.1; -.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0019031; C:viral envelope; IEA.
```

```

DR GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR002548; Alpha_E1_glycop.
DR InterPro: IPR000936; Alpha_E2_glycop.
DR InterPro: IPR002533; Alpha_E3_glycop.
DR InterPro: IPR000336; Flavi_glycoprote.
DR Pfam: PF01589; Alpha_E1_glycop; 1.
DR Pfam: PF00943; Alpha_E2_glycop; 1.
DR Pfam: PF01563; Alpha_E3_glycop; 1.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 559
SQ SEQUENCE 559 AA; 62045 MW; 6646981CBB17B50 CRC64;

Query Match 48.5%; Score 48; DB 2; Length 559;
Best Local Similarity 50.0%; Pred. No. 61;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SSDGLMNNNQTOLE 16
Db 487 SLDLHMNNQOMFWIQ 502

RESULT 24
ID Q9WCGO PRELIMINARY; PRT; 559 AA.
AC Q9WCGO;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Structural polyprotein (Fragment).
OS Venezuelan equine encephalitis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus; VEEV complex.
OX NCBI_TaxID=11036;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL23/69;
RX MEDLINE=99394886; PubMed=10466974;
RA Weaver S.C., Pfeiffer M., Marriott K., Kang W., Kinney R.M.;
RT "Genetic evidence for the origins of Venezuelan equine encephalitis
RT virus subtype IAB outbreaks."
RL Am. J. Trop. Med. Hyg. 60:441-448(1999).
DR EMBL: AF093105; AAD37004.1; -.
DR GO:0019028; C:viral envelope; IEA.
DR GO:0019031; C:viral capsid; IEA.
DR GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR002548; Alpha_E1_glycop.
DR InterPro: IPR000936; Alpha_E2_glycop.
DR InterPro: IPR002533; Alpha_E3_glycop.
DR InterPro: IPR000336; Flavi_glycoprote.
DR Pfam: PF01589; Alpha_E1_glycop; 1.
DR Pfam: PF00943; Alpha_E2_glycop; 1.
DR Pfam: PF01563; Alpha_E3_glycop; 1.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 559
SQ SEQUENCE 559 AA; 61882 MW; E03AD37DAE6AB54 CRC64;

Query Match 48.5%; Score 48; DB 2; Length 559;
Best Local Similarity 50.0%; Pred. No. 61;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SSDGLMNNNQTOLE 16
Db 487 SLDLHMNNQOMFWIQ 502

RESULT 25
ID Q98755 PRELIMINARY; PRT; 559 AA.
AC Q98755;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)

```

```

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Structural polyprotein (Fragment).
OS Venezuelan equine encephalitis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus; VEEV complex.
OX NCBI_TaxID=11036;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V178;
RX MEDLINE=96332169; PubMed=8709783;
RA Weaver S.C., Salas R., Rico-Hesse R., Ludwig G.V., Oberste M.S.,
RA Boshell J., Fesh R.B.;
RT "Re-emergence of epidemic Venezuelan equine encephalomyelitis in South
RT America. VEE Study Group."
RL Lancet 348:436-440(1996).
DR EMBL: U55341; AAC58459.1; -.
DR GO:0019028; C:viral capsid; IEA.
DR GO:0019031; C:viral envelope; IEA.
DR GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR002548; Alpha_E1_glycop.
DR InterPro: IPR000936; Alpha_E2_glycop.
DR InterPro: IPR002533; Alpha_E3_glycop.
DR InterPro: IPR000336; Flavi_glycoprote.
DR Pfam: PF01589; Alpha_E1_glycop; 1.
DR Pfam: PF00943; Alpha_E2_glycop; 1.
DR Pfam: PF01563; Alpha_E3_glycop; 1.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 537
FT CHAIN 538 >559 protein E2.
FT NON_TER 559 protein E1.
SQ SEQUENCE 559 AA; 61711 MW; DEA2018E20C30308 CRC64;

Query Match 48.5%; Score 48; DB 2; Length 559;
Best Local Similarity 50.0%; Pred. No. 61;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SSDGLMNNNQTOLE 16
Db 487 SLDLHMNNQOMFWIQ 502

RESULT 26
ID Q8VSJ7 PRELIMINARY; PRT; 559 AA.
AC Q8VSJ7;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Structural polyprotein (Fragment).
OS Venezuelan equine encephalitis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus; VEEV complex.
OX NCBI_TaxID=11036;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V178;
RX MEDLINE=96332169; PubMed=8709783;
RA Weaver S.C., Salas R., Rico-Hesse R., Ludwig G.V., Oberste M.S.,
RA Boshell J., Fesh R.B.;
RT "Re-emergence of epidemic Venezuelan equine encephalomyelitis in South
RT America. VEE Study Group."
RL Lancet 348:436-440(1996).
DR EMBL: U55341; AAC58459.1; -.
DR GO:0019028; C:viral capsid; IEA.
DR GO:0019031; C:viral envelope; IEA.
DR GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR002548; Alpha_E1_glycop.
DR InterPro: IPR000936; Alpha_E2_glycop.
DR InterPro: IPR002533; Alpha_E3_glycop.
DR InterPro: IPR000336; Flavi_glycoprote.
DR Pfam: PF01589; Alpha_E1_glycop; 1.
DR Pfam: PF00943; Alpha_E2_glycop; 1.
DR Pfam: PF01563; Alpha_E3_glycop; 1.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 559
SQ SEQUENCE 559 AA; 61823 MW; 9D14BFB740780862 CRC64;

```

```

Query Match      48.5%; Score 48; DB 2; Length 559;
Best Local Similarity 50.0%; Pred. No. 61;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      1 SSDGLMNNQOTQFLFLE 16
Db      487 SLDLHNNNQOMFWIQ 502

RESULT 27
ID      08V5J8      PRELIMINARY;      PRT;      559 AA.
AC      08V5J8;
DT      01-MAR-2002 (TREMBLrel. 20, Created)
DT      01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE      Structural polyprotein (Fragment).
OS      Venezuelan equine encephalitis virus.
OC      Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
CX      Alphavirus; VEEV complex.
NC      NCB1_TaxID=11036;
RN      SEQUENCE FROM N.A.
RP      Weaver S.C.;
RL      Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF348335; AAL57186.1;
DR      GO; GO:0019028; C:Viral capsid; IEA.
DR      GO; GO:0019031; C:Viral envelope; IEA.
DR      GO; GO:0005198; F:Structural molecule activity; IEA.
DR      InterPro; IPR002548; Alpha_E1_glycop.
DR      InterPro; IPR000936; Alpha_E2_glycop.
DR      InterPro; IPR002533; Alpha_E3_glycop.
DR      InterPro; IPR000336; Flav1_glycoprote.
DR      Pfam; PF01589; Alpha_E1_glycop; 1.
DR      Pfam; PF00943; Alpha_E2_glycop; 1.
DR      Pfam; PF01563; Alpha_E3_glycop; 1.
KW      Polypeptid.
FT      NON_TER      1      1
FT      NON_TER      559      559
SQ      SEQUENCE 559 AA; 61823 MW; 9014BFB740780862 CRC64;

Query Match      48.5%; Score 48; DB 2; Length 559;
Best Local Similarity 50.0%; Pred. No. 61;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      1 SSDGLMNNQOTQFLFLE 16
Db      487 SLDLHNNNQOMFWIQ 502

RESULT 28
ID      POLS_EEVV8      STANDARD;      PRT;      1254 AA.
AC      P05674;
DT      01-NOV-1988 (Rel. 09, Created)
DT      01-NOV-1988 (Rel. 09, Last sequence update)
DT      01-OCT-2004 (Rel. 45, Last annotation update)
DE      Structural polyprotein (P130) [Contains: Coat protein C (EC 3.4.21.-)
DE      (Capsid protein C); Spike glycoprotein E3; Spike glycoprotein E2; 6
DE      kDa peptide; Spike glycoprotein E1].
OS      Venezuelan equine encephalitis virus (strain TC-83).
OC      Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
CX      Alphavirus.
NC      NCB1_TaxID=11037;
RN      SEQUENCE FROM N.A.
RP      MEDLINE=66306669; PubMed=3755750;
RX      Johnson B.J.B., Kinney R.M., Kost C.L., Trent D.W.;
RT      "Molecular determinants of alphavirus neurovirulence: nucleotide and
RT      deduced protein sequence changes during attenuation of Venezuelan
RT      equine encephalitis virus.";
RL      J. Gen. Virol. 67:1951-1960(1986).

```

```

CC      -!- FUNCTION: The capsid protein is an auto-protease.
CC      -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC      -!- MISCELLANEOUS: The 6 kDa polypeptide probably serves as the signal
CC      sequence for the membrane glycoprotein E1, which is the viral
CC      hemagglutinin.
CC      -!- SIMILARITY: The protease belongs to peptidase family S3.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
CC      EMBL; X04368; CA27883.1; -.
CC      PIR; A27871; VHWVVE.
CC      PDB; 1EP5; X-ray; A/B/C=119-275.
CC      PDB; 1EP6; X-ray; A/B/C=119-275.
CC      MEROPS; S03.001; -.
CC      InterPro; IPR002548; Alpha_E1_glycop.
CC      InterPro; IPR000936; Alpha_E2_glycop.
CC      InterPro; IPR002533; Alpha_E3_glycop.
CC      InterPro; IPR000336; Flav1_glycoprote.
CC      InterPro; IPR000930; Peptidase S3.
CC      InterPro; IPR009003; Pept_Ser_Cys.
CC      Pfam; PF00944; Alpha_core; 1.
CC      Pfam; PF01589; Alpha_E1_glycop; 1.
CC      Pfam; PF00943; Alpha_E2_glycop; 1.
CC      Pfam; PF01563; Alpha_E3_glycop; 1.
CC      PRINTS; PR00798; TOGAVIRIN.
CC      3D-structure: Coat protein; Glycoprotein; Hydrolase; Polypeptid;
CC      Serine protease; Transmembrane.
CC      KW      CHAIN      1      275
CC      FT      CHAIN      276      334      Spike glycoprotein E3.
CC      FT      CHAIN      335      757      Spike glycoprotein E2.
CC      FT      CHAIN      758      812      6 kDa peptide.
CC      FT      CHAIN      813      1254      Spike glycoprotein E1.
CC      FT      ACT_SITE      152      152      Charge relay system (By similarity).
CC      FT      ACT_SITE      158      158      Charge relay system (By similarity).
CC      FT      ACT_SITE      226      226      Charge relay system (By similarity).
CC      FT      TRANSMEM      702      718      Potential.
CC      FT      TRANSMEM      774      790      Potential.
CC      FT      TRANSMEM      792      808      Potential.
CC      FT      TRANSMEM      1231      1247      Potential.
CC      FT      CARBOHYD      286      286      N-linked (GlcNAc...) (Potential).
CC      FT      CARBOHYD      546      546      N-linked (GlcNAc...) (Potential).
CC      FT      CARBOHYD      652      652      N-linked (GlcNAc...) (Potential).
CC      FT      CARBOHYD      946      946      N-linked (GlcNAc...) (Potential).
CC      SQ      SEQUENCE 1254 AA; 138485 MW; 7615698519A529F6 CRC64;

Query Match      48.5%; Score 48; DB 1; Length 1254;
Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      1 SSDGLMNNQOTQFLFLE 16
Db      763 SLDLHNNNQOMFWIQ 778

RESULT 29
ID      POLS_EEVVE      STANDARD;      PRT;      1254 AA.
AC      P36330; Q66582; Q66583; Q66584; Q66585; Q66586;
DT      01-JUN-1994 (Rel. 29, Created)
DT      01-JUN-1994 (Rel. 29, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Structural polyprotein (P130) [Contains: Coat protein C (EC 3.4.21.-)
DE      (Capsid protein C); Spike glycoprotein E3; Spike glycoprotein E2; 6
DE      kDa peptide; Spike glycoprotein E1].
OS      Venezuelan equine encephalitis virus (strain Everglades Fe3-7c).
OC      Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
CX      Alphavirus.

```

```

OX NCB1_TaxID=36383;
RN [1]
RN SEQUENCE FROM N.A.
RX Sneider J.M., Kinney R.M., Tsuchiya K.R., Trent D.W.;
RT "Molecular evidence that epizootic Venezuelan equine encephalitis
RT (VEE) I-A-B viruses are not evolutionary derivatives of enzootic VEE
RT subtype I-E or II viruses.";
RL J. Gen. Virol. 74:519-523 (1993).
CC -1- FUNCTION: The capsid protein is an auto-protease.
CC -1- PPM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -1- MISCELLANEOUS: The 6 kDa polypeptide probably serves as the signal
CC sequence for the membrane glycoprotein E1, which is the viral
CC hemagglutinin.
CC -1- SIMILARITY: The protease belongs to peptidase family S3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L04598; AAA42984.1; -;
DR EMBL; L04598; AAA42983.1; ALT TERM.
DR EMBL; L04598; AAA42985.1; ALT SEQ.
DR EMBL; L04598; AAA42986.1; ALT SEQ.
DR EMBL; L04598; AAA42987.1; ALT SEQ.
DR EMBL; L04598; AAA42988.1; ALT_INIT.
DR PIR; J01978; J01978.
DR HSSP; P05674; 1EP5.
DR INTERPRO; IPR002548; Alpha_E1_glycop.
DR INTERPRO; IPR000936; Alpha_E2_glycop.
DR INTERPRO; IPR002533; Alpha_E3_glycop.
DR INTERPRO; IPR000336; Flav1_glycoprote.
DR INTERPRO; IPR009030; Peptidase_S3.
DR Pfam; PF00944; Alpha_core1.1.
DR Pfam; PF00589; Alpha_E1_glycop; 1.
DR Pfam; PF00943; Alpha_E2_glycop; 1.
DR Pfam; PF01563; Alpha_E3_glycop; 1.
DR PRINTS; PR00798; TOGAVIRIN.
KM Coat protein; Glycoprotein; Hydroxylase; Polyprotein; Serine protease;
KW Transmembrane.
FT CHAIN 1 274 Coat protein C.
FT CHAIN 275 333 Spike glycoprotein E3.
FT CHAIN 334 756 Spike glycoprotein E2.
FT CHAIN 757 812 6 kDa peptide.
FT CHAIN 813 1254 Spike glycoprotein E1.
FT ACT_SITE 151 151 Charge relay system (By similarity).
FT ACT_SITE 157 157 Charge relay system (By similarity).
FT ACT_SITE 225 225 Charge relay system (By similarity).
FT TRANSMEM 701 721 Potential.
FT TRANSMEM 794 813 Potential.
FT TRANSMEM 1231 1248 Potential.
FT CARBOHYD 47 47 N-linked (G1CNAc. . .) (Potential).
FT CARBOHYD 285 285 N-linked (G1CNAc. . .) (Potential).
FT CARBOHYD 545 545 N-linked (G1CNAc. . .) (Potential).
FT CARBOHYD 651 651 N-linked (G1CNAc. . .) (Potential).
FT CARBOHYD 946 946 N-linked (G1CNAc. . .) (Potential).
SQ SEQUENCE 1254 AA; 138337 MW; 29DDEF37F9E92C4B CRC64;
Query Match 48.5%; Score 48; DB 1; Length 1254;
Best Local Similarity 57.9%; Pred. No. 1.5e+02;
Matches 11; Conservative 1; Mismatches 3; Indels 4; Gaps 1;
OY 1 SSDGLANNNQ-----TOLFL 15
Db 762 SLDIHANNNQMFWTQLLI 780

```

```

RESULT 30
ID POLS_EEVT STANDARD; PRT; 1254 AA.
AC P09592; 088691; 088692; 088693; 088694; 088695;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Structural polyprotein (P130) [contains: Coat protein C (EC 3.4.21.-)
DE (Capsid protein C); Spike glycoprotein E3; Spike glycoprotein E2; 6
DE kDa peptide; Spike glycoprotein E1].
OS Venezuelan equine encephalitis virus (strain Trinidad donkey).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus.
OX NCB1_TaxID=11038;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=86263392; PubMed=3088830;
RA Kinney R.M., Johnson B.J.B., Welch J.B., Tsuchiya K.R., Trent D.W.;
RT "Nucleotide sequence of the 26 S mRNA of the virulent Trinidad donkey
RT strain of Venezuelan equine encephalitis virus and deduced sequence of
RT the encoded structural proteins.";
RL Virology 152:400-413 (1986).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=89243175; PubMed=2524126;
RA Kinney R.M., Johnson B.J.B., Welch J.B., Tsuchiya K.R., Trent D.W.;
RT "The full-length nucleotide sequences of the virulent Trinidad donkey
RT strain of Venezuelan equine encephalitis virus and its attenuated
RT vaccine derivative strain TC-83.";
RL Virology 170:19-30 (1989).
CC -1- FUNCTION: The capsid protein is an auto-protease.
CC -1- PPM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -1- MISCELLANEOUS: The 6 kDa polypeptide probably serves as the signal
CC sequence for the membrane glycoprotein E1, which is the viral
CC hemagglutinin.
CC -1- SIMILARITY: The protease belongs to peptidase family S3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14937; AAA42997.1; -;
DR EMBL; J04332; AAB02519.1; -;
DR PIR; B31467; VHWVYT.
DR HSSP; P05674; 1EP5.
DR INTERPRO; IPR002548; Alpha_E1_glycop.
DR INTERPRO; IPR000936; Alpha_E2_glycop.
DR INTERPRO; IPR002533; Alpha_E3_glycop.
DR INTERPRO; IPR000336; Flav1_glycoprote.
DR INTERPRO; IPR009030; Peptidase_S3.
DR Pfam; PF00944; Alpha_core1.1.
DR Pfam; PF00589; Alpha_E1_glycop; 1.
DR Pfam; PF00943; Alpha_E2_glycop; 1.
DR Pfam; PF01563; Alpha_E3_glycop; 1.
DR PRINTS; PR00798; TOGAVIRIN.
KM Coat protein; Glycoprotein; Hydroxylase; Polyprotein; Serine protease;
KW Transmembrane.
FT CHAIN 1 275 Coat protein C.
FT CHAIN 276 334 Spike glycoprotein E3.
FT CHAIN 335 757 Spike glycoprotein E2.
FT CHAIN 758 812 6 kDa peptide.
FT CHAIN 813 1254 Spike glycoprotein E1.
FT ACT_SITE 152 152 Charge relay system (By similarity).
FT ACT_SITE 158 158 Charge relay system (By similarity).
FT ACT_SITE 226 226 Charge relay system (By similarity).
FT TRANSMEM 702 718 Potential.
FT TRANSMEM 774 790 Potential.

```

```
FT TRANSMEM 792 808 Potential.
FT TRANSMEM 1231 1247 Potential.
FT CARBOHYD 286 286 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 546 546 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 652 652 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 946 946 N-linked (GlcNAc... ) (Potential).
FT CONFLECT 811 811 P -> G (in Ref. 2).
SQ SEQUENCE 1254 AA; 138390 MW; 1C0F62BE1909664B CRC64;

Query Match 48.5%; Score 48; DB 1; Length 1254;
Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 SSDGLANNNOQTOLFL 16
Db 763 SLDLHNNNOQMFWIQ 778

RESULT 31
ID Q77VY8 PRELIMINARY; PRT; 1254 AA.
AC Q77VY8;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DE Structural polypeptide.
OS Venezuelan equine encephalitis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus; VEEV complex.
OX NCBI_TaxID=11036;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Everglaides Fe3-7c;
RX MEDLINE=99101297; PubMed=9886206;
RA Kinney R.M., Pfeiffer M., Tsuchiya K.R., Chang G.J., Roehrig J.T.;
RT "Nucleotide sequences of the 26S mRNAs of the viruses defining the
RT Venezuelan equine encephalitis antigenic complex.";
RL Am. J. Trop. Med. Hyg. 59:952-964(1998).
DR EMBL; AF075251; AAD14551.1;
DR InterPro; IPR002548; Alpha_E1_glycop.
DR InterPro; IPR000936; Alpha_E2_glycop.
DR InterPro; IPR002533; Alpha_E3_glycop.
DR InterPro; IPR000336; Flavi_glycoprote.
DR InterPro; IPR000930; Peptidase_S3.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00944; Alpha_core; 1.
DR Pfam; PF01589; Alpha_E1_glycop; 1.
DR Pfam; PF00943; Alpha_E2_glycop; 1.
DR Pfam; PF01563; Alpha_E3_glycop; 1.
DR PRINTS; PR00798; TOGAVIRIN.
KM Polypeptide.
FT CHAIN 813 1254 envelope glycoprotein 1.
FT CHAIN 1 274 nucleocapsid protein.
FT CHAIN 275 333 polypeptide cleavage product E3.
FT CHAIN 334 756 envelope glycoprotein 2.
FT CHAIN 757 812 6K polypeptide cleavage product.
SQ SEQUENCE 1254 AA; 138337 MW; 29DBEF37F9B92C4B CRC64;

Query Match 48.5%; Score 48; DB 2; Length 1254;
Best Local Similarity 57.9%; Pred. No. 1.5e+02;
Matches 11; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

QY 1 SSDGLANNNOQTOLFL 15
Db 762 SLDLHNNNOQMFWTOLLI 780

RESULT 32
ID Q9YKC4 PRELIMINARY; PRT; 1254 AA.
AC Q9YKC4;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
```

```
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Structural polypeptide.
OS Venezuelan equine encephalitis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus; VEEV complex.
OX NCBI_TaxID=11036;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Cabasso Caar 508;
RX MEDLINE=99101297; PubMed=9886206;
RA Kinney R.M., Pfeiffer M., Tsuchiya K.R., Chang G.J., Roehrig J.T.;
RT "Nucleotide sequences of the 26S mRNAs of the viruses defining the
RT Venezuelan equine encephalitis antigenic complex.";
RL Am. J. Trop. Med. Hyg. 59:952-964(1998).
DR EMBL; AF075259; AAD14567.1;
DR HSSP; P06674; IEPS.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002548; Alpha_E1_glycop.
DR InterPro; IPR000936; Alpha_E2_glycop.
DR InterPro; IPR002533; Alpha_E3_glycop.
DR InterPro; IPR000336; Flavi_glycoprote.
DR InterPro; IPR000930; Peptidase_S3.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00944; Alpha_core; 1.
DR Pfam; PF01589; Alpha_E1_glycop; 1.
DR Pfam; PF00943; Alpha_E2_glycop; 1.
DR Pfam; PF01563; Alpha_E3_glycop; 1.
DR PRINTS; PR00798; TOGAVIRIN.
KM Polypeptide.
FT CHAIN 813 1254 envelope glycoprotein 1.
FT CHAIN 1 275 nucleocapsid protein.
FT CHAIN 276 334 polypeptide cleavage product E3.
FT CHAIN 335 756 envelope glycoprotein 2.
FT CHAIN 757 812 6K polypeptide cleavage product.
SQ SEQUENCE 1254 AA; 138639 MW; 3ED95B5ABE786CFE CRC64;

Query Match 48.5%; Score 48; DB 2; Length 1254;
Best Local Similarity 57.1%; Pred. No. 1.5e+02;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 DGLMNNNOQTOLFL 16
Db 764 DGLMNNNOQMFWLQ 777

RESULT 33
ID AAD14551 PRELIMINARY; PRT; 1254 AA.
AC AAD14551;
DT 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE Structural polypeptide.
OS Venezuelan equine encephalitis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus.
OX NCBI_TaxID=11036;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Everglaides Fe3-7c;
RX MEDLINE=99101297; PubMed=9886206;
RA Kinney R.M., Pfeiffer M., Tsuchiya K.R., Chang G.J., Roehrig J.T.;
RT "Nucleotide sequences of the 26S mRNAs of the viruses defining the
RT Venezuelan equine encephalitis antigenic complex.";
RL Am. J. Trop. Med. Hyg. 59:952-964(1998).
DR EMBL; AF075251; AAD14551.1;
DR CHAIN 1 274 NUCLEOCAPSID PROTEIN.
FT CHAIN 275 333 POLYPEPTIDE CLEAVAGE PRODUCT E3.
```

FT CHAIN 334 756 ENVELOPE GLYCOPROTEIN 2.
 FT CHAIN 757 812 6K POLYPROTEIN CLEAVAGE PRODUCT.
 FT CHAIN 813 1254 ENVELOPE GLYCOPROTEIN 1.
 SQ SEQUENCE 1254 AA; 138337 MW; 29DDEF37F9BE92C4B CRC64;

Query Match 48.5%; Score 48; DB 2; Length 1254;
 Best Local Similarity 57.9%; Pred. No. 1.5e+02;
 Matches 11; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

Oy 1 SSDGLMNNNO---TOLFL 15
 Db 762 SLDHLNNNOQMFWTLLI 780

RESULT 34
 POLS_EEYV3 STANDARD; PRT: 1255 AA.
 AC P36332;

DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Structural polyprotein (P130) [Contains: Coat protein C (EC 3.4.21.-)
 DE (capsid protein C); Spike glycoprotein E3; Spike glycoprotein E2; 6
 DE kDa peptide; Spike glycoprotein E1].
 OS Venezuelan equine encephalitis virus (strain 3880).
 OC Alphaviruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
 NC NCB1_TaxID=36382;

RN SEQUENCE FROM N.A.
 RA MEDLINE=93079859; PubMed=1448915;
 RX Kinney R.M., Tsuchiya K.R., Snelder J.M., Trent D.W.;
 RT "Genetic evidence that epizootic Venezuelan equine encephalitis (VEE)
 RT viruses may have evolved from enzootic VEE subtype I-D virus.";
 RL Virology 191:569-580(1992).

CC -1- FUNCTION: The capsid protein is an auto-protease.
 CC -1- PFM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -1- MISCELLANEOUS: The 6 kDa polypeptide probably serves as the signal
 CC sequence for the membrane glycoprotein E1, which is the viral
 CC haemagglutinin.
 CC -1- SIMILARITY: The protease belongs to peptidase family S3.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; L00930; AAC19325.1; -.
 CC PIR; D44213; D44213.
 CC HSSP; P05674; 1EP5.
 CC MEROPS; S03.001; -.
 DR InterPro; IPR002548; Alpha_E1_glycop.
 DR InterPro; IPR000936; Alpha_E2_glycop.
 DR InterPro; IPR002533; Alpha_E3_glycop.
 DR InterPro; IPR000336; Flavi_glycoproteB.
 DR InterPro; IPR000930; Peptidase_S3.
 DR InterPro; IPR009003; Pept_Ser_Cys.
 DR Pfam; PF00944; Alpha_core; 1.
 DR Pfam; PF01589; Alpha_E1_glycop; 1.
 DR Pfam; PF00943; Alpha_E2_glycop; 1.
 DR Pfam; PF01563; Alpha_E3_glycop; 1.
 DR PRINTS; PR00798; TOGAVIRIN.

KW Coat protein; Glycoprotein; Hydroxylase; Polypeptide; Serine protease;
 KW Transmembrane.

FT CHAIN 1 275 Coat protein C.
 FT CHAIN 276 334 Spike glycoprotein E3.
 FT CHAIN 335 757 Spike glycoprotein E2.
 FT CHAIN 758 813 6 kDa peptide.
 FT CHAIN 814 1255 Spike glycoprotein E1.
 FT ACT_SITE 152 Charge relay system (By similarity).

FT ACT_SITE 158 158 Charge relay system (By similarity).
 FT ACT_SITE 226 226 Charge relay system (By similarity).
 FT TRANSMEM 702 722 Potential.
 FT TRANSMEM 795 814 Potential.
 FT TRANSMEM 1232 1249 Potential.
 FT CARBOHYD 47 47 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 286 286 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 652 652 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 947 947 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 1255 AA; 138297 MW; 7D30E17CAECC310 CRC64;

Query Match 48.5%; Score 48; DB 1; Length 1255;
 Best Local Similarity 50.0%; Pred. No. 1.5e+02;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Oy 1 SSDGLMNNNOTOLFL 16
 Db 763 SLDHLNNNOQMFWIQ 778

RESULT 35
 POLS_EEYV3 STANDARD; PRT: 1255 AA.
 AC P36332;

DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Structural polyprotein (P130) [Contains: Coat protein C (EC 3.4.21.-)
 DE (capsid protein C); Spike glycoprotein E3; Spike glycoprotein E2; 6
 DE kDa peptide; Spike glycoprotein E1].
 OS Venezuelan equine encephalitis virus (strain P676).
 OC Alphaviruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
 NC NCB1_TaxID=36385;

RN SEQUENCE FROM N.A.
 RA MEDLINE=93079859; PubMed=1448915;
 RX Kinney R.M., Tsuchiya K.R., Snelder J.M., Trent D.W.;
 RT "Genetic evidence that epizootic Venezuelan equine encephalitis (VEE)
 RT viruses may have evolved from enzootic VEE subtype I-D virus.";
 RL Virology 191:569-580(1992).

CC -1- FUNCTION: The capsid protein is an auto-protease.
 CC -1- PFM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -1- MISCELLANEOUS: The 6 kDa polypeptide probably serves as the signal
 CC sequence for the membrane glycoprotein E1, which is the viral
 CC haemagglutinin.
 CC -1- SIMILARITY: The protease belongs to peptidase family S3.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; L04653; AAC19319.1; -.
 CC PIR; B44213; B44213.
 CC HSSP; P05674; 1EP5.
 CC MEROPS; S03.001; -.
 DR InterPro; IPR002548; Alpha_E1_glycop.
 DR InterPro; IPR000936; Alpha_E2_glycop.
 DR InterPro; IPR002533; Alpha_E3_glycop.
 DR InterPro; IPR000336; Flavi_glycoproteB.
 DR InterPro; IPR000930; Peptidase_S3.
 DR InterPro; IPR009003; Pept_Ser_Cys.
 DR Pfam; PF00944; Alpha_core; 1.
 DR Pfam; PF01589; Alpha_E1_glycop; 1.
 DR Pfam; PF00943; Alpha_E2_glycop; 1.
 DR Pfam; PF01563; Alpha_E3_glycop; 1.
 DR PRINTS; PR00798; TOGAVIRIN.

KW Coat protein; Glycoprotein; Hydroxylase; Polypeptide; Serine protease;
 KW Transmembrane.

DR Pfam; PF00943; Alpha_E2_glycop; 1.
 DR Pfam; PF01563; Alpha_E3_glycop; 1.
 DR PRINTS; PR00798; TOGAVIRIN.
 KM Polypeptidn.
 SQ SEQUENCE 1255 AA; 138564 MW; 9229511B3907A399 CRC64;

Query Match 48.5%; Score 48; DB 2; Length 1255;
 Best Local Similarity 50.0%; Pred. No. 1.5e+02;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SSDGLMNNNOTQFLFLE 16
 Db 763 SLDLMLNNNOQMFWIQ 778

RESULT 38
 ID 036300 PRELIMINARY; PRT; 1255 AA.

AC 036300;
 DT 01-JAN-1998 (TREMBlrel. 05; Created)
 DT 01-MAY-1999 (TREMBlrel. 10; Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26; Last annotation update)
 DE Structural polypeptidn.
 GN Name=nsP4;
 OS Venezuelan equine encephalitis virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
 OC Alphavirus; VEEV complex.
 OX NCBI_Taxid=11036;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Beck/Mycoff;
 RX MEDLINE=97404681; PubMed=9261393;
 RA Smith J.F., Kang W., Sweeney M.S., Brault A.C., Rico-Hesse R., Schmura S.M., Powers A.M., Oberste M.S., Weaver S.C.;
 RT "Repeated emergence of epidemic/epizootic Venezuelan equine encephalitis from a single genotype of enzootic subtype ID virus."
 RL J. Virol. 71:6697-6705(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Beck/Mycoff;
 RA Powers A.M., Oberste M.S., Brault A.C., Rico-Hesse R., Schmura S.M., Smith J.F., Kang W., Sweeney W.P., Weaver S.C.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF004471; AAC36374.1; -.
 DR HSSP; P05674; IEP5.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR002548; Alpha_E1_glycop.
 DR InterPro; IPR000936; Alpha_E2_glycop.
 DR InterPro; IPR002533; Alpha_E3_glycop.
 DR InterPro; IPR000336; Flavi_glycoprote.
 DR InterPro; IPR000930; Peptidase_S3.
 DR InterPro; IPR009003; Pept_Ser_Cys.
 DR Pfam; PF00944; Alpha_core1_1.
 DR Pfam; PF01569; Alpha_E1_glycop; 1.
 DR Pfam; PF00943; Alpha_E2_glycop; 1.
 DR Pfam; PF01563; Alpha_E3_glycop; 1.
 DR PRINTS; PR00798; TOGAVIRIN.
 KM Polypeptidn.
 FT CHAIN 538 1255 envelope glycoprotein E1.
 FT CHAIN 1 537 envelope glycoprotein E2.
 SQ SEQUENCE 1255 AA; 138476 MW; 2C1BCB6F196EC751 CRC64;

Query Match 48.5%; Score 48; DB 2; Length 1255;
 Best Local Similarity 50.0%; Pred. No. 1.5e+02;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 Qy 1 SSDGLMNNNOTQFLFLE 16
 Db 763 SLDLMLNNNOQMFWIQ 778

Db 763 SLDLMLNNNOQMFWIQ 778

RESULT 39

ID 090164 PRELIMINARY; PRT; 1255 AA.

AC 090164;
 DT 01-NOV-1998 (TREMBlrel. 08; Created)
 DT 01-NOV-1998 (TREMBlrel. 08; Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26; Last annotation update)
 DE Structural polypeptidn.
 OS Venezuelan equine encephalitis virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
 OC Alphavirus; VEEV complex.
 OX NCBI_Taxid=11036;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=71-180;
 RX MEDLINE=89243175; PubMed=2524126;
 RA Kinney R.M., Johnson B.J., Welch J.B., Tsuchiya K.R., Trent D.W.;
 RT "The full-length nucleotide sequences of the virulent Trinidad donkey strain of Venezuelan equine encephalitis virus and its attenuated vaccine derivative strain TC-83";
 RL Virology 170:19-30(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=71-180;
 RX MEDLINE=93107872; PubMed=1469368;
 RA Kinney R.M., Tsuchiya K.R., Sneider J.M., Trent D.W.;
 RT "Molecular evidence for the origin of the widespread Venezuelan equine encephalitis epizootic of 1969 to 1972."
 RL J. Gen. Virol. 73:0-0(0).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=71-180;
 RA Kinney R.M., Trent D.W.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF069903; AAC24034.1; -.
 DR HSSP; P05674; IEP5.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR002548; Alpha_E1_glycop.
 DR InterPro; IPR000936; Alpha_E2_glycop.
 DR InterPro; IPR002533; Alpha_E3_glycop.
 DR InterPro; IPR000336; Flavi_glycoprote.
 DR InterPro; IPR009003; Peptidase_S3.
 DR InterPro; IPR009003; Pept_Ser_Cys.
 DR Pfam; PF00944; Alpha_core1_1.
 DR Pfam; PF01569; Alpha_E1_glycop; 1.
 DR Pfam; PF00943; Alpha_E2_glycop; 1.
 DR Pfam; PF01563; Alpha_E3_glycop; 1.
 DR PRINTS; PR00798; TOGAVIRIN.
 KM Polypeptidn.
 FT CHAIN 814 1255 E1.
 FT CHAIN 1 275 capsid protein.
 FT CHAIN 276 334 E3.
 FT CHAIN 335 757 E2.
 FT CHAIN 758 813 E2.
 SQ SEQUENCE 1255 AA; 138458 MW; 8A733DC7501404C2 CRC64;

Query Match 48.5%; Score 48; DB 2; Length 1255;
 Best Local Similarity 50.0%; Pred. No. 1.5e+02;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SSDGLMNNNOTQFLFLE 16
 Db 763 SLDLMLNNNOQMFWIQ 778

RESULT 40

```

091628      PRELIMINARY;      PRT; 1255 AA.
ID 091628
AC 091628;
DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Structural glycoprotein.
GN Name:sp4;
OS Venezuelan equine encephalitis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus; VEEV complex.
OX NCBI_TaxID=11036;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6921;
RA Smith J.F., Kang W., Oberste M.S., Brault A.C., Rico-Hesse R., Schmura S.M.,
RA Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
RL EMBL; AF004470; AAC36375.1; -.
DR HSBP; P05674; IEP5.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteinolysis and peptidolysis; IEA.
DR InterPro; IPR002548; Alpha_E1_glycop.
DR InterPro; IPR000936; Alpha_E2_glycop.
DR InterPro; IPR002533; Alpha_E3_glycop.
DR InterPro; IPR000336; Flavi_glycoprote.
DR InterPro; IPR000930; Peptidase_S3.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00944; Alpha_core1_1.
DR Pfam; PF01589; Alpha_E1_glycop; 1.
DR Pfam; PF00943; Alpha_E2_glycop; 1.
DR Pfam; PF01563; Alpha_E3_glycop; 1.
DR PRINTS; PR00798; TOGAVIRIN.
KM Polypeptide.
FT CHAIN 538 1255 envelope glycoprotein E1.
FT CHAIN 1 537 envelope glycoprotein E2.
SQ SEQUENCE 1255 AA; 138365 MW; F08FA923B9C0BA2 CRC64;

Query Match 48.5%; Score 48; DB 2; Length 1255;
Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 SSDGLANNNTQLFLR 16
Db 763 SLDLHNNNQMFMIQ 778

```

Search completed: December 30, 2004, 16:13:13
 Job time : 196 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: December 30, 2004, 15:55:44 ; Search time 154 Seconds

(without alignments)
41.929 Million cell updates/sec

Title: US-10-718-321-1

Perfect score: 99

Sequence: 1 SSDGLMNNQTLFLEHS 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729239 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : A_Geneseq_23Sep04.*

1: Geneseq1980s.*

2: Geneseq1990s.*

3: Geneseq2000s.*

4: Geneseq2001s.*

5: Geneseq2002s.*

6: Geneseq2003as.*

7: Geneseq2003bs.*

8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	99	100.0	18	AAO26687	Aao26687 Human KIM
2	99	100.0	18	AAO26673	Aao26673 Monoclonal
3	99	100.0	81	AAO26678	Aao26678 Human KIM
4	99	100.0	263	AAO26629	Aao26629 Human NOV
5	99	100.0	263	AAO26622	Aao26622 Plasmid P
6	99	100.0	263	AAO26596	Aao26596 Human NOV
7	99	100.0	334	AAO26336	Aao26336 Human kid
8	99	100.0	334	AAO26679	Aao26679 KIM-1 rel
9	99	100.0	339	AAO26618	Aao26618 Plasmid P
10	99	100.0	339	AAO26592	Aao26592 Human NOV
11	99	100.0	359	AAO26582	Aao26582 Human can
12	99	100.0	359	AAO26439	Aao26439 Amino aci
13	99	100.0	359	AAO26438	Aao26438 Amino aci
14	99	100.0	359	AAO26417	Aao26417 Human bla
15	99	100.0	359	AAO26680	Aao26680 KIM-1 rel
16	99	100.0	359	AAO26594	Aao26594 Human NOV
17	99	100.0	359	AAO26584	Aao26584 Cancer/an
18	99	100.0	364	AAO26443	Aao26443 Amino aci
19	99	100.0	364	AAO26440	Aao26440 Amino aci
20	99	100.0	365	AAO26441	Aao26441 Amino aci
21	99	100.0	365	AAO26441	Aao26441 Amino aci
22	99	100.0	365	AAO26441	Aao26441 Amino aci
23	99	100.0	365	AAO26441	Aao26441 Amino aci
24	99	100.0	365	AAO26441	Aao26441 Amino aci
25	99	100.0	365	AAO26441	Aao26441 Amino aci

26	46	508	5	ABG93049	Abg93049 S. cerevi
27	46	508	6	ABR52836	AbR52836 Protein s
28	46	508	6	ADK62270	AdK62270 Disease t
29	46	545	6	ABG71366	AbG71366 Potato pl
30	46	545	8	ADJ49214	AdJ49214 Oil-aseoc
31	45	514	8	ABR57094	AbR57094 MLHR comp
32	45	514	6	ABR57095	AbR57095 MLHR comp
33	45	244	4	AAO42107	Aao42107 Human pol
34	45	244	4	AAO42108	Aao42108 Human pol
35	45	265	3	AAO42109	Aao42109 Human pol
36	45	301	7	ADD46355	Add46355 Rat Prote
37	45	301	7	ADD46357	Add46357 Rat Prote
38	45	309	4	AAO42470	Aao42470 Human pro
39	45	309	5	ABR97591	AbR97591 Novel hum
40	45	511	4	AAO40322	Aao40322 Human pol
41	45	588	4	AAO40321	Aao40321 Human pol
42	45	588	4	AAO40321	Aao40321 Human pol
43	45	607	5	AAO16525	Aao16525 Inulin-1
44	45	687	6	ABU20598	Abu20598 Protein e
45	45	1464	2	AAO44192	Aao44192 Rat NMDA
46	45	1464	5	AAO44192	Aao44192 Rat NMDA
47	45	1464	5	AAO44192	Aao44192 Rat NMDA
48	45	1464	5	AAO44192	Aao44192 Rat NMDA
49	45	1464	5	AAO44192	Aao44192 Rat NMDA
50	45	1464	5	AAO44192	Aao44192 Rat NMDA
51	45	1464	5	AAO44192	Aao44192 Rat NMDA
52	45	1464	5	AAO44192	Aao44192 Rat NMDA
53	45	1464	5	AAO44192	Aao44192 Rat NMDA
54	45	1464	5	AAO44192	Aao44192 Rat NMDA
55	45	1464	5	AAO44192	Aao44192 Rat NMDA
56	45	1464	5	AAO44192	Aao44192 Rat NMDA
57	45	1464	5	AAO44192	Aao44192 Rat NMDA
58	45	1464	5	AAO44192	Aao44192 Rat NMDA
59	45	1464	5	AAO44192	Aao44192 Rat NMDA
60	45	1464	5	AAO44192	Aao44192 Rat NMDA
61	45	1464	5	AAO44192	Aao44192 Rat NMDA
62	45	1464	5	AAO44192	Aao44192 Rat NMDA
63	45	1464	5	AAO44192	Aao44192 Rat NMDA
64	45	1464	5	AAO44192	Aao44192 Rat NMDA
65	45	1464	5	AAO44192	Aao44192 Rat NMDA
66	45	1464	5	AAO44192	Aao44192 Rat NMDA
67	45	1464	5	AAO44192	Aao44192 Rat NMDA
68	45	1464	5	AAO44192	Aao44192 Rat NMDA
69	45	1464	5	AAO44192	Aao44192 Rat NMDA
70	45	1464	5	AAO44192	Aao44192 Rat NMDA
71	45	1464	5	AAO44192	Aao44192 Rat NMDA
72	45	1464	5	AAO44192	Aao44192 Rat NMDA
73	45	1464	5	AAO44192	Aao44192 Rat NMDA
74	45	1464	5	AAO44192	Aao44192 Rat NMDA
75	45	1464	5	AAO44192	Aao44192 Rat NMDA
76	45	1464	5	AAO44192	Aao44192 Rat NMDA
77	45	1464	5	AAO44192	Aao44192 Rat NMDA
78	45	1464	5	AAO44192	Aao44192 Rat NMDA
79	45	1464	5	AAO44192	Aao44192 Rat NMDA
80	45	1464	5	AAO44192	Aao44192 Rat NMDA
81	45	1464	5	AAO44192	Aao44192 Rat NMDA
82	45	1464	5	AAO44192	Aao44192 Rat NMDA
83	45	1464	5	AAO44192	Aao44192 Rat NMDA
84	45	1464	5	AAO44192	Aao44192 Rat NMDA
85	45	1464	5	AAO44192	Aao44192 Rat NMDA
86	45	1464	5	AAO44192	Aao44192 Rat NMDA
87	45	1464	5	AAO44192	Aao44192 Rat NMDA
88	45	1464	5	AAO44192	Aao44192 Rat NMDA
89	45	1464	5	AAO44192	Aao44192 Rat NMDA
90	45	1464	5	AAO44192	Aao44192 Rat NMDA
91	45	1464	5	AAO44192	Aao44192 Rat NMDA
92	45	1464	5	AAO44192	Aao44192 Rat NMDA
93	45	1464	5	AAO44192	Aao44192 Rat NMDA
94	45	1464	5	AAO44192	Aao44192 Rat NMDA
95	45	1464	5	AAO44192	Aao44192 Rat NMDA
96	45	1464	5	AAO44192	Aao44192 Rat NMDA
97	45	1464	5	AAO44192	Aao44192 Rat NMDA
98	45	1464	5	AAO44192	Aao44192 Rat NMDA

99	41	41.4	326	7	ADE36761	Adc36761 Rhesus ro
100	41	41.4	399	4	ABG19483	Abg19483 Novel hum
101	41	41.4	487	4	AAO0109	AAO0109 Beta-amy1
102	41	41.4	494	5	ABE53478	ABE53478 Lactococc
103	41	41.4	540	8	ADJ48637	Adj48637 O11-aseoc
104	41	41.4	542	8	ADJ48616	Adj48616 O11-aseoc
105	41	41.4	569	6	ABU35742	Abu35742 Protein e
106	41	41.4	602	4	ABG07448	Abg07448 Novel hum
107	41	41.4	624	5	ABP73354	Abp73354 Candida a
108	41	41.4	721	7	ADJ70091	Adj70091 Human hea
109	41	41.4	790	4	ABU53243	Abu53243 Human sig
110	41	41.4	790	5	AAE23390	AAE23390 Human int
111	41	41.4	790	7	ADB75262	ADB75262 Prostate
112	41	41.4	871	3	AAAB41558	AAAB41558 Human ORF
113	41	41.4	871	4	AAAB3936	AAAB3936 Human pol
114	41	41.4	925	2	AAE21606	AAE21606 G6 amy1as
115	41	41.4	947	6	ABU11694	Abu11694 Human MDD
116	41	41.4	947	6	ABU11527	Abu11527 Human MDD
117	41	41.4	956	3	AAV94752	AAV94752 Non-malco
118	41	41.4	1281	5	ABG32157	Abg32157 N-methyl
119	41	41.4	1350	4	AAU02944	AAU02944 Angiotens
120	41	41.4	1455	2	AAE23824	AAE23824 Murine ma
121	41	41.4	1464	2	AAE66039	AAE66039 Human N-m
122	41	41.4	1464	2	AAE55529	AAE55529 Human NMD
123	41	41.4	1464	2	AAE80970	AAE80970 Human exc
124	41	41.4	1464	2	AAE85576	AAE85576 Human N-m
125	41	41.4	1464	3	AAV56113	AAV56113 Human NMD
126	41	41.4	1464	3	AAE26215	AAE26215 Human N-m
127	41	41.4	1464	5	AAE56507	AAE56507 Human NMD
128	41	41.4	1464	5	AAO18063	AAO18063 Human NMD
129	41	41.4	1464	5	ABG32158	Abg32158 N-methyl
130	41	41.4	1464	6	ABU65878	Abu65878 Human N-m
131	41	41.4	1464	6	ABU61423	Abu61423 Human N-m
132	41	41.4	1464	6	ABU03575	Abu03575 Human exp
133	41	41.4	1464	6	ABU03578	Abu03578 Human exp
134	41	41.4	1464	6	ABU03574	Abu03574 Human exp
135	41	41.4	1464	6	ABU03577	Abu03577 Human exp
136	41	41.4	1464	6	ABU03572	Abu03572 Human exp
137	41	41.4	1464	6	ABU03579	Abu03579 Human exp
138	41	41.4	1464	6	ABU03571	Abu03571 Human exp
139	41	41.4	1464	6	ABU03576	Abu03576 Human exp
140	41	41.4	1464	6	ABU03580	Abu03580 Human exp
141	41	41.4	1464	6	ABU03570	Abu03570 Human exp
142	41	41.4	1464	6	ABU57660	Abu57660 Human N-m
143	41	41.4	1464	6	ABO43070	ABO43070 Human NMD
144	41	41.4	1464	7	AAE39234	AAE39234 Human NMD
145	41	41.4	1970	4	ABE64827	ABE64827 Drosoph11
146	40	40.4	126	8	ADL05841	ADL05841 M. catarr
147	40	40.4	141	3	AAO1614	AAO1614 Human sec
148	40	40.4	149	4	ABG05062	ABG05062 Novel hum
149	40	40.4	161	2	AAE93597	AAE93597 Almond N-
150	40	40.4	161	2	AAE93595	AAE93595 Almond N-

ALIGNMENTS

RESULT 1

AAO26687 ID AAO26687 standard; peptide; 18 AA.

XX AAO26687;

DT 20-MAR-2003 (first entry)

XX Human KIM-1 mucin domain related peptide, SEQ ID No 15.

XX Cytostatic; gene therapy; antibody; antigen; antigen-binding;

XX proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;

XX renal disease; injury; renal cancer.

XX Synthetic.

PN	WO200298920-A1.	
XX	12-DEC-2002.	
PD	31-MAY-2002; 2002WO-US017402.	
XX	01-JUN-2001; 2001US-0295449P.	
XX	04-JUN-2001; 2001US-0295907P.	
PR	(BIOJ) BIOGEN INC.	
PA	(GEHO) GEN HOSPITAL CORP.	
XX	Bailly V, Bonventre J;	
PI	WPI; 2003-156845/15.	
XX		
DR		
XX		
FT	New antibody, antibody derivative or antigen-binding polypeptide that	
PT	inhibits proteolytic release of a soluble kidney injury molecule-1	
PT	polypeptide, useful for treating or preventing renal disease or injury,	
PT	e.g. renal cancer.	
XX		
PS	Disclosure; Fig 1A; 42pp; English.	
XX		
CC	The invention relates to a novel antibody, antibody derivative or antigen	
CC	-binding polypeptide that inhibits proteolytic release of a soluble	
CC	kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.	
CC	The antibody, antibody derivative or antigen-binding polypeptide is	
CC	useful for treating or preventing renal disease or injury, e.g. renal	
CC	cancer. The antibody is also useful for inhibiting shedding of the KIM-1	
CC	peptide. This sequence represents a human KIM-1 mucin domain related	
XX		
SQ	Sequence 18 AA;	

Query Match 100.0%; Score 99; DB 6; Length 18;
 Best Local Similarity 100.0%; Pred. No. 4.7e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 SSDGLMNNOTOLFLEHS 18
 DB 1 SSDGLMNNOTOLFLEHS 18

RESULT 2

AAO26673 ID AAO26673 standard; peptide; 18 AA.

XX AAO26673;

DT 20-MAR-2003 (first entry)

XX Monoclonal antibody binding epitope, SEQ ID No 1.

XX Cytostatic; gene therapy; antibody; antigen; antigen-binding;

XX proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;

XX renal disease; injury; renal cancer; binding epitope.

XX Homo sapiens.

XX WO200298920-A1.

PD 12-DEC-2002.

PF 31-MAY-2002; 2002WO-US017402.

XX 01-JUN-2001; 2001US-0295449P.

XX 04-JUN-2001; 2001US-0295907P.

XX (BIOJ) BIOGEN INC.

XX (GEHO) GEN HOSPITAL CORP.

XX Bailly V, Bonventre J;

XX

DR WPI; 2003-156845/15.
XX New antibody, antibody derivative or antigen-binding polypeptide that
PT inhibits proteolytic release of a soluble kidney injury molecule-1
PT polypeptide, useful for treating or preventing renal disease or injury,
PT e.g. renal cancer.
XX
PS Claim 4; Page 25; 42pp; English.
XX
CC The invention relates to a novel antibody, antibody derivative or antigen
CC -binding polypeptide that inhibits proteolytic release of a soluble
CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
CC The antibody, antibody derivative or antigen-binding polypeptide is
CC useful for treating or preventing renal disease or injury, e.g. renal
CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
CC polypeptide. This sequence represents a binding epitope of the invention
XX
SQ Sequence 18 AA;
Query Match 100.0%; Score 99; DB 6; Length 18;
Best Local Similarity 100.0%; Pred. No. 4,7e-08; Mismatches 0; Gaps 0;
Matches 18; Conservative 0; Indels 0; Gaps 0;
QY 1 SSDGLMNNNOTQFLFHS 18
1 SSDGLMNNNOTQFLFHS 18
Db 1 SSDGLMNNNOTQFLFHS 18
RESULT 3
AAO26678
ID AAO26678 standard; protein; 81 AA.
XX
XX AAO26678;
AC
XX
XX 20-MAR-2003 (first entry)
DT
XX
XX Human KIM-1 mucin domain protein, SEQ ID NO 6.
DE
XX
XX Cytostatic; gene therapy; antibody; antigen; antigen-binding;
KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
KW renal disease; injury; renal cancer; human.
XX
XX Homo sapiens.
OS
XX
XX WO200298920-A1.
PN
XX
XX 12-DEC-2002.
PD
XX
XX 31-MAY-2002; 2002WO-US017402.
PF
XX
XX 01-JUN-2001; 2001US-0295449P.
PR
XX
XX 04-JUN-2001; 2001US-0295907P.
PR
XX
XX (BIOJ) BIOGEN INC.
PA (GEHO) GEN HOSPITAL CORP.
XX
XX
XX Batilly V, Bonventre J;
PI
XX
XX WPI; 2003-156845/15.
DR
XX
XX New antibody, antibody derivative or antigen-binding polypeptide that
PT inhibits proteolytic release of a soluble kidney injury molecule-1
PT polypeptide, useful for treating or preventing renal disease or injury,
PT e.g. renal cancer.
XX
XX
XX Disclosure; Fig 1A; 42pp; English.
PS
XX
XX The invention relates to a novel antibody, antibody derivative or antigen
CC -binding polypeptide that inhibits proteolytic release of a soluble
CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
CC The antibody, antibody derivative or antigen-binding polypeptide is
CC useful for treating or preventing renal disease or injury, e.g. renal
CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1

CC polypeptide. This sequence represents a human KIM-1 mucin domain protein
CC of the invention
XX
XX Sequence 81 AA;
SQ
QY 1 SSDGLMNNNOTQFLFHS 18
1 SSDGLMNNNOTQFLFHS 18
Db 55 SSDGLMNNNOTQFLFHS 72
Query Match 100.0%; Score 99; DB 6; Length 81;
Best Local Similarity 100.0%; Pred. No. 2,3e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
RESULT 4
ADE36628
ID ADE36628 standard; protein; 263 AA.
XX
XX ADE36629;
AC
XX
XX 29-JAN-2004 (first entry)
DT
XX
XX Human NOVX protein SEQ ID NO:39.
DE
XX
XX human; NOVX; cytosolic; antiinflammatory; gene therapy; vaccine;
KW renal cancer; inflammation; tissue typing.
XX
XX Homo sapiens.
OS
XX
XX WO2003080856-A2.
PN
XX
XX 02-OCT-2003.
PD
XX
XX 19-MAR-2003; 2003NO-US008490.
PF
XX
XX 19-MAR-2002; 2002US-0365491P.
PR
XX
XX 13-SEP-2002; 2002US-0410618P.
PR
XX
XX (CURA-) CURAGEN CORP.
PA
XX
XX Lepley DM, Rieger DK, Tse K, Rastelli L, Smithson G, Mesri M;
PI Ooi CE, Anderson DM, Guo X, Giot L, Starling G;
XX
XX WPI; 2003-876927/81.
DR
XX
XX New polypeptide, useful for preparing a composition for treating or
PT preventing a pathology associated with NOVX polypeptide e.g. renal cancer
PT or inflammation, or for tissue typing.
PT
XX
XX Example 9; SEQ ID NO 39; 239pp; English.
PS
XX
XX The present invention describes an isolated human NOVX polypeptide, where
CC X is 1a to 1d or 2a to 2h, or its mature form, a sequence that is at
CC least 95 % identical to 1c, or a sequence comprising one or more
CC conservative substitutions in the amino acid sequence. The NOVX
CC polypeptide, and nucleic acid sequence encoding it, has cytosolic and
CC antiinflammatory activities, and can be used in gene therapy, and in
CC vaccines. The NOVX polypeptide is useful for preparing a composition for
CC treating or preventing a pathology associated with NOVX polypeptide e.g.
CC renal cancer or inflammation, or for tissue typing. The present sequence
CC represents a protein sequence which is used in an example from the
CC present invention.
XX
XX
XX Sequence 263 AA;
SQ
Query Match 100.0%; Score 99; DB 7; Length 263;
Best Local Similarity 100.0%; Pred. No. 8,2e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SSDGLMNNNOTQFLFHS 18
1 SSDGLMNNNOTQFLFHS 18
Db 244 SSDGLMNNNOTQFLFHS 261

```

RESULT 5
ADE36622
ID ADE36622 standard; protein; 263 AA.
XX
AC ADE36622;
XX
DT 29-JAN-2004 (first entry)
XX
DE Plasmid PCR2.1-CG57008-02-S84a 13A protein insert SEQ ID NO:32.
XX
KW human; NOVX; cytostatic; antiinflammatory; gene therapy; vaccine;
KM renal cancer; inflammation; tissue typing.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO2003080856-A2.
XX
PD 02-OCT-2003.
XX
PF 19-MAR-2003; 2003WO-US008490.
XX
PR 19-MAR-2002; 2002US-0355491P.
PR 13-SEP-2002; 2002US-0410618P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Lepley DM, Rieger DK, Tse K, Rastelli L, Smithson G, Mesri M;
PI Ool CE, Anderson DM, Guo X, Giot L, Starling G;
XX
DR WPI; 2003-876927/81.
DR N-PSDB; ADE36621.
XX
PT New polypeptide, useful for preparing a composition for treating or
PT preventing a pathology associated with NOVX polypeptide e.g. renal cancer
PT or inflammation, or for tissue typing.
XX
PS Example 2; SEQ ID NO 32; 239pp; English.
XX
SQ The present invention describes an isolated human NOVX polypeptide, where
XX
CC X is 1a to 1d or 2a to 2h, or its mature form, a sequence that is at
CC least 95 % identical to 1f, or a sequence comprising one or more
CC conservative substitutions in the amino acid sequence. The NOVX
CC polypeptide, and nucleic acid sequence encoding it, has cytostatic and
CC antiinflammatory activities, and can be used in gene therapy, and in
CC vaccines. The NOVX polypeptide is useful for preparing a composition for
CC treating or preventing a pathology associated with NOVX polypeptide e.g.
CC renal cancer or inflammation, or for tissue typing. The present sequence
CC represents a PCR2.1-CG57008-02-S841 13A protein insert, which is used in
CC an example from the present invention.
XX
SQ Sequence 263 AA;
XX
Query Match 100.0%; Score 99; DB 7; Length 263;
Best Local Similarity 100.0%; Pred. No. 8.2e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SSDGLMNNNOTQLFLEHS 18
DB 244 SSDGLMNNNOTQLFLEHS 261
XX
RESULT 6
ADE36596
ID ADE36596 standard; protein; 263 AA.
XX
AC ADE36596;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human NOVXc protein SEQ ID NO:6.
XX

```

```

KW human; NOVX; cytostatic; antiinflammatory; gene therapy; vaccine;
KM renal cancer; inflammation; tissue typing.
XX
OS Homo sapiens.
XX
PN WO2003080856-A2.
XX
PD 02-OCT-2003.
XX
PF 19-MAR-2003; 2003WO-US008490.
XX
PR 19-MAR-2002; 2002US-0355491P.
PR 13-SEP-2002; 2002US-0410618P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Lepley DM, Rieger DK, Tse K, Rastelli L, Smithson G, Mesri M;
PI Ool CE, Anderson DM, Guo X, Giot L, Starling G;
XX
DR WPI; 2003-876927/81.
DR N-PSDB; ADE36595.
XX
PT New polypeptide, useful for preparing a composition for treating or
PT preventing a pathology associated with NOVX polypeptide e.g. renal cancer
PT or inflammation, or for tissue typing.
XX
PS Claim 1; SEQ ID NO 6; 239pp; English.
XX
SQ The present invention describes an isolated human NOVX polypeptide, where
XX
CC X is 1a to 1d or 2a to 2h, or its mature form, a sequence that is at
CC least 95 % identical to 1f, or a sequence comprising one or more
CC conservative substitutions in the amino acid sequence. The NOVX
CC polypeptide, and nucleic acid sequence encoding it, has cytostatic and
CC antiinflammatory activities, and can be used in gene therapy, and in
CC vaccines. The NOVX polypeptide is useful for preparing a composition for
CC treating or preventing a pathology associated with NOVX polypeptide e.g.
CC renal cancer or inflammation, or for tissue typing. The present sequence
CC represents human NOVXc from the present invention.
XX
SQ Sequence 263 AA;
XX
Query Match 100.0%; Score 99; DB 7; Length 263;
Best Local Similarity 100.0%; Pred. No. 8.2e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SSDGLMNNNOTQLFLEHS 18
DB 244 SSDGLMNNNOTQLFLEHS 261
XX
RESULT 7
AAW38336
ID AAW38336 standard; protein; 334 AA.
XX
AC AAW38336;
XX
DT 21-MAY-1998 (first entry)
XX
DE Human kidney injury related molecule (KIM).
XX
KW kidney injury related molecule; KIM; human; renal disease; injury;
KM nephritis; tissue regeneration; therapy; monoclonal antibody.
XX
OS Homo sapiens.
XX
PN WO9744460-A1.
XX
PD 27-NOV-1997.
XX
PF 23-MAY-1997; 97WO-US009303.
XX
PR 24-MAY-1996; 96US-0018228P.
PR 23-AUG-1996; 96US-0023442P.
XX

```

XX (BIOJ) BIOGEN INC.
XX PA Sanicola-Nadel M, Bonventre JV, Hession CA, Ichimura T, Wei H,
XX PI Cate RL;
XX DR WPI; 1998-018514/02.
XX PT DNA encoding kidney injury related molecule - which is upregulated in
XX PT injured or regenerating tissue, useful to promote growth of new tissue
XX PT and survival of damaged tissue.
XX PS Claim 9; Page 46-47; 68pp; English.
XX
XX This protein, designated kidney injury related molecule (KIM), is up-
XX regulated in injured or regenerating tissue. Its amino acid sequence was
XX deduced from a clone (see AAT96035) obtained from a human embryonic liver
XX library. A 572-amino acid (see AAW38335) and a 307-amino acid rat KIM
XX (see AAW38334) are also claimed. Recombinant KIM polypeptides can be
XX expressed in prokaryotic and eukaryotic host cells using a claimed
XX process. Soluble variants fused to a toxin, imageable compound or
XX radionuclide, and IgG fusion proteins are also claimed. KIM, or an
XX agonist, can be used to treat renal disease and to promote the growth of new
XX tissue or the survival of damaged tissue, generally in conditions where
XX the binding of specific ligand to KIM stimulates cell growth, maintains
XX cellular differentiation or reduces apoptosis, e.g. in cases of renal
XX failure, hepatitis, kidney transplants, toxic or hypoxic injury. A
XX monoclonal antibody specific for KIM can be used to treat renal disease,
XX e.g. where binding of KIM to ligand results in neoplasia, loss of
XX cellular function, susceptibility to apoptosis or promotion of
XX inflammation, deliver imaging agents to KIM expressing cells in vivo or
XX in vitro and measure KIM concentration by immunoassay.
XX Damage/regeneration of renal cells can be determined by measuring KIM,
XX particularly to diagnose or monitor the progress of disease or therapy.
XX KIM-expressing tumour cells can be inhibited by treatment with a fusion
XX protein comprising KIM ligand or Mab with a toxin or radionuclide, and
XX tumour cells that express KIM ligand can be inhibited with similarly
XX tagged KIM or anti-KIM ligand antibody
XX
XX Sequence 334 AA:
SQ
Query Match 100.0%; Score 99; DB 2; Length 334;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 18; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SSDGLMNNNOTQLFLEHS 18
Db 264 SSDGLMNNNOTQLFLEHS 281
RESULT 8
AAO26679
ID AAO26679 standard; protein; 334 AA.
XX
XX AAO26679;
XX
XX 20-MAR-2003 (first entry)
XX
XX KIM-1 related protein, SEQ ID NO 7.
XX
XX Cytostatic; gene therapy; antibody; antigen; antigen-binding;
XX proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
XX renal disease; injury; renal cancer; human.
XX Homo sapiens.
XX OS
XX MO200298920-A1.
XX PN
XX 12-DEC-2002.
XX PD
XX 31-MAY-2002; 2002WO-US017402.
XX PF
XX 01-JUN-2001; 2001US-0295449P.
XX PR

PR 04-JUN-2001; 2001US-0295907P.
XX
XX (BIOJ) BIOGEN INC.
XX PA (GENO) GEN HOSPITAL CORP.
XX PI Bailly V, Bonventre J;
XX DR WPI; 2003-156845/15.
XX
XX New antibody, antibody derivative or antigen-binding polypeptide that
XX PT inhibits proteolytic release of a soluble kidney injury molecule-1
XX PT polypeptide, useful for treating or preventing renal disease or injury,
XX PT e.g. renal cancer.
XX
XX Disclosure; Page 39-40; 42pp; English.
XX
XX The invention relates to a novel antibody, antibody derivative or antigen
XX -binding polypeptide that inhibits proteolytic release of a soluble
XX kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
XX The antibody, antibody derivative or antigen-binding polypeptide is
XX useful for treating or preventing renal disease or injury, e.g. renal
XX cancer. The antibody is also useful for inhibiting shedding of the KIM-1
XX polypeptide. This sequence represents a protein relating to the KIM-1
XX protein of the invention
XX
XX Sequence 334 AA:
SQ
Query Match 100.0%; Score 99; DB 6; Length 334;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 18; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SSDGLMNNNOTQLFLEHS 18
Db 264 SSDGLMNNNOTQLFLEHS 281
RESULT 9
ADE36618
ID ADE36618 standard; protein; 339 AA.
XX
XX ADE36618;
XX
XX 29-JAN-2004 (first entry)
XX
XX Plasmid PCR2.1-CG57008-03-S843 15B protein insert SEQ ID NO:28.
XX DE
XX human; NOX; cyostatic; antiinflammatory; gene therapy; vaccine;
XX KW renal cancer; inflammation; tissue typing.
XX
XX Synthetic.
XX OS
XX Homo sapiens.
XX MO2003080856-A2.
XX PN
XX 02-OCT-2003.
XX PD
XX 19-MAR-2003; 2003WO-US008490.
XX PP
XX 19-MAR-2002; 2002US-0365491P.
XX PR 13-SEP-2002; 2002US-0410618P.
XX PR
XX (CURA-) CURAGEN CORP.
XX PA
XX Lepley DW, Rieger DK, Tse K, Rastelli L, Smithson G, Neerl M;
XX PI Ooi CR, Anderson DW, Guo X, Giot L, Starling G;
XX DR WPI; 2003-876927/81.
XX DR N-PSDB; ADE36617.
XX
XX New polypeptide, useful for preparing a composition for treating or
XX PT preventing a pathology associated with NOX polypeptide e.g. renal cancer
XX PT or inflammation, or for tissue typing.
XX PR

PS Example 1; SEQ ID NO 28; 239pp; English.
XX
CC The present invention describes an isolated human NOVX polypeptide, where
CC X is 1a to 1d or 2a to 2h, or its mature form, a sequence that is at
CC least 95 % identical to it, or a sequence comprising one or more
CC conservative substitutions in the amino acid sequence. The NOVX
CC polypeptide, and nucleic acid sequence encoding it, has cytostatic and
CC antiinflammatory activities, and can be used in gene therapy, and in
CC vaccines. The NOVX polypeptide is useful for preparing a composition for
CC treating or preventing a pathology associated with NOVX polypeptide e.g.
CC renal cancer or inflammation, or for tissue typing. The present sequence
CC represents a PCR2.1-CG57008-03-S843 15B protein insert, which is used in
CC an example from the present invention.
XX
SQ Sequence 339 AA;

Query Match 100.0%; Score 99; DB 7; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSDGLMNNQTLFLEHS 18
|||
Db 244 SSDGLMNNQTLFLEHS 261

RESULT 10
ADE36592
ID ADE36592 standard; protein; 339 AA.
XX
AC ADE36592;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human NOV1a protein SEQ ID NO:2.
XX
KW human; NOVX; cytoskeletal; antiinflammatory; gene therapy; vaccine;
KW renal cancer; inflammation; tissue typing.
XX
OS Homo sapiens.
XX
PN WO2003080856-A2.
XX
PD 02-OCT-2003.
XX
PF 19-MAR-2003; 2003WO-US008490.
XX
PR 19-MAR-2002; 2002US-0365491P.
PR 13-SEP-2002; 2002US-0410618P.
PA (CURA-) CURAGEN CORP.
PI Lepley DM, Rieger DK, Tse K, Rascelli L, Smithson G, Mesri M;
PI Ooi CE, Anderson DW, Guo X, Giot L, Starling G;
XX
DR WPI; 2003-876927/81.
DR N-PSDB; ADE36591.
XX
PT New polypeptide, useful for preparing a composition for treating or
PT preventing a pathology associated with NOVX polypeptide e.g. renal cancer
PT or inflammation, or for tissue typing.
XX
PS Claim 1; SEQ ID NO 2; 239pp; English.
XX
CC The present invention describes an isolated human NOVX polypeptide, where
CC X is 1a to 1d or 2a to 2h, or its mature form, a sequence that is at
CC least 95 % identical to it, or a sequence comprising one or more
CC conservative substitutions in the amino acid sequence. The NOVX
CC polypeptide, and nucleic acid sequence encoding it, has cytostatic and
CC antiinflammatory activities, and can be used in gene therapy, and in
CC vaccines. The NOVX polypeptide is useful for preparing a composition for
CC treating or preventing a pathology associated with NOVX polypeptide e.g.
CC renal cancer or inflammation, or for tissue typing. The present sequence
CC represents human NOV1a from the present invention.

XX
SQ Sequence 339 AA;

Query Match 100.0%; Score 99; DB 7; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSDGLMNNQTLFLEHS 18
|||
Db 244 SSDGLMNNQTLFLEHS 261

RESULT 11
ABR58582
ID ABR58582 standard; protein; 359 AA.
XX
AC ABR58582;
XX
DT 09-JUL-2003 (first entry)
XX
DE Human cancer related protein SEQ ID NO:239.
XX
KW Human; cancer; diagnosis; screening; modulator; leukaemia; ischaemia;
KW heart disease; atherosclerosis; endometriosis.
XX
OS Homo sapiens.
XX
PN WO2003025138-A2.
XX
PD 27-MAR-2003.
XX
PF 17-SEP-2002; 2002WO-US029560.
XX
PR 17-SEP-2001; 2001US-0323469P.
PR 20-SEP-2001; 2001US-0323887P.
PR 13-NOV-2001; 2001US-0350667P.
PR 08-FEB-2002; 2002US-0355145P.
PR 08-FEB-2002; 2002US-0355257P.
PR 12-APR-2002; 2002US-0372246P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE;
PI Zlotnik A;
XX
DR WPI; 2003-354600/33.
DR N-PSDB; ACC72709.
XX
XX New genes that are up-regulated or down-regulated in cancers, useful as
PT markers for diagnosing e.g. cancer, ischemia or heart diseases, or as
PT therapeutic targets for screening drugs for treating these diseases.
XX
PS Claim 12; Page 742; 767pp; English.
XX
CC The present invention describes an isolated nucleic acid molecule, which
CC comprises the sequence of any of the genes that are up-regulated or down-
CC regulated in specific cancers (e.g. about 1031 genes up-regulated in
CC acute lymphocytic leukemia) . ACC72641 to ACC72860 represent cancer
CC related gene nucleotide sequences which encode the proteins given in
CC ABR58582 to ABR58709. Also described: (1) determining the presence or
CC absence of a pathological cell in a patient; (2) an expression vector
CC comprising a nucleic acid molecule described above; (3) a host cell
CC comprising the vector; (4) an isolated polypeptide, which is encoded by
CC the nucleic acid; (5) an antibody that specifically binds the polypeptide
CC of (4); (6) specifically targeting a compound to a pathological cell in a
CC patient by administering to the patient the antibody above; and (7) a
CC drug screening assay. The nucleic acid is useful as diagnostic markers or
CC therapeutic targets. In particular, the nucleic acid is useful for
CC diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,
CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,
CC pancreas, prostate, skin and uterus), wounds, ischaemia, heart diseases,
CC atherosclerosis and endometriosis. The nucleic acid is also useful in
CC drug screening, particularly for identifying agents for treating these

CC pathologies
XX
SQ Sequence 359 AA;

Query Match 100.0%; Score 99; DB 6; Length 359;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SSDGLMNNQOTLFLEHS 18
Db 264 SSDGLMNNQOTLFLEHS 281

RESULT 12

ABP70439
ID ABP70439 standard; protein; 359 AA.

XX
AC ABP70439;

XX
DT 22-APR-2003 (first entry)

XX
DE Amino acid sequence of human TIM-1 allele 1.

XX
KW T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;
KW TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;
KW myelodysplastic syndrome; airway hyperreactivity; cancer; asthma;
KW allergic T cell response; autoimmune disease.

XX
OS Homo sapiens.

XX
PN WC2003002722-A2.

XX
PD 09-JAN-2003.

XX
PF 01-JUL-2002; 2002WO-US020890.

XX
PR 29-JUN-2001; 2001US-0302344P.

XX
PA (STRD) UNIV LELAND STANFORD JUNIOR.

XX
PI McIntire JJ, Dekruyff RH, Umetsu DT, Freeman GJ, Kuchroo V;

XX
DR WPI; 2003-210268/20.

XX
DR N-PSDB; ABZ68333.

XX
PT New nucleic acid comprising a mammalian T cell immunoglobulin domain and
PT Mucin domain gene sequences, useful for treating cancer or asthma,
PT allergy, eczema or autoimmune disease.

XX
PS Claim 10; Page 82; 94pp; English.

XX
CC The present sequence is a human T cell immunoglobulin domain and mucin
CC domain-1 (TIM-1) polypeptide. The specification describes TIM-1, TIM-2,
CC TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with
CC conserved IGV and mucin domains. The locus comprising the TIM family is
CC genetically associated with immune dysfunction, including asthma. The TIM
CC gene family is located within a region of human chromosome 5 that is
CC commonly deleted in malignancies and myelodysplastic syndrome. Variants
CC of TIM-1 and TIM-3 are associated with susceptibility to airway
CC hyperreactivity and allergic T cell responses, and other variants
CC associated with protection against these responses. T cells express TIM
CC proteins, which critically regulate CD4 T cell differentiation. Th1 cells
CC preferentially express TIM-3, while Th2 cells preferentially express TIM-
CC 1. TIM polypeptides and polynucleotides are useful for treating cancer,
CC asthma, allergies, eczema or autoimmune diseases

XX
SQ Sequence 359 AA;

Query Match 100.0%; Score 99; DB 6; Length 359;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SSDGLMNNQOTLFLEHS 18

Db 264 SSDGLMNNQOTLFLEHS 281

RESULT 13

ABP70438
ID ABP70438 standard; protein; 359 AA.

XX
AC ABP70438;

XX
DT 22-APR-2003 (first entry)

XX
DE Amino acid sequence of human TIM-1 allele 1.

XX
KW T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;
KW TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;
KW myelodysplastic syndrome; airway hyperreactivity; cancer; asthma;
KW allergic T cell response; autoimmune disease.

XX
OS Homo sapiens.

XX
PN WC2003002722-A2.

XX
PD 09-JAN-2003.

XX
PF 01-JUL-2002; 2002WO-US020890.

XX
PR 29-JUN-2001; 2001US-0302344P.

XX
PA (STRD) UNIV LELAND STANFORD JUNIOR.

XX
PI McIntire JJ, Dekruyff RH, Umetsu DT, Freeman GJ, Kuchroo V;

XX
DR WPI; 2003-210268/20.

XX
DR N-PSDB; ABZ68332.

XX
PT New nucleic acid comprising a mammalian T cell immunoglobulin domain and
PT Mucin domain gene sequences, useful for treating cancer or asthma,
PT allergy, eczema or autoimmune disease.

XX
PS Claim 10; Page 80-81; 94pp; English.

XX
CC The present sequence is a human T cell immunoglobulin domain and mucin
CC domain-1 (TIM-1) polypeptide. The specification describes TIM-1, TIM-2,
CC TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with
CC conserved IGV and mucin domains. The locus comprising the TIM family is
CC genetically associated with immune dysfunction, including asthma. The TIM
CC gene family is located within a region of human chromosome 5 that is
CC commonly deleted in malignancies and myelodysplastic syndrome. Variants
CC of TIM-1 and TIM-3 are associated with susceptibility to airway
CC hyperreactivity and allergic T cell responses, and other variants
CC associated with protection against these responses. T cells express TIM
CC proteins, which critically regulate CD4 T cell differentiation. Th1 cells
CC preferentially express TIM-3, while Th2 cells preferentially express TIM-
CC 1. TIM polypeptides and polynucleotides are useful for treating cancer,
CC asthma, allergies, eczema or autoimmune diseases

XX
SQ Sequence 359 AA;

Query Match 100.0%; Score 99; DB 6; Length 359;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SSDGLMNNQOTLFLEHS 18
Db 264 SSDGLMNNQOTLFLEHS 281

RESULT 14

ABR48174
ID ABR48174 standard; protein; 359 AA.

XX
AC ABR48174;

XX 12-JUN-2003 (first entry)
XX Human bladder cancer associated protein sequence SEQ ID NO:64.
DE Human bladder cancer; cytostatic; gene therapy; vaccine.
XX Homo sapiens.
OS
XX MO2003003906-A2.
PN
XX 16-JAN-2003.
PD
XX 03-JUN-2002; 2002WO-US021338.
XX
XX 03-JUN-2001; 2001US-0302814P.
PR 03-AUG-2001; 2001US-0310099P.
PR 08-NOV-2001; 2001US-0343705P.
PR 13-NOV-2001; 2001US-0350666P.
PR 12-APR-2002; 2002US-0372246P.
XX
XX (EOSB-) EOS BIOTECHNOLOGY INC.
PA
XX Mack DH, Aziz N;
PI
XX WPI; 2003-201532/19.
DR N-PSDB; ACC50985.
DR
XX
XX Detecting a bladder cancer-associated transcript in a cell from a
PT patient, comprises contacting a biological sample from the patient with a
PT bladder cancer-associated polynucleotide or antibody.
XX
XX Claim 10; Page 252; 307pp; English.
XX
XX The present invention describes a method for detecting a bladder cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with a polynucleotide
CC that selectively hybridizes to a sequence that is 80 % identical to a
CC table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
CC encode the human bladder cancer-associated proteins given in ABR48146 to
CC ABR48242). Bladder cancer-associated sequences from the present invention
CC have cytostatic activities, and can be used in antisense gene therapy and
CC in vaccine production. The method can be used for detecting a bladder
CC cancer-associated transcript in a cell from a patient. The method is
CC useful in diagnosing or treating bladder cancer and in screening for
CC compounds that modulate bladder cancer, such as hormones or antibodies.
CC The nucleic acid molecules from the present invention may be used in
CC various screening and diagnostic methods, and for gene therapy, vaccine
CC and/or antisense/inhibition applications
XX
XX Sequence 359 AA;
SQ
Query Match 100.0%; Score 99; DB 6; Length 359;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SSDGLMNNNOTQFLFHS 18
Db 264 SSDGLMNNNOTQFLFHS 281
RESULT 15
AAO26680
ID AAO26680 standard; protein; 359 AA.
XX
XX AAO26680;
AC
XX 20-MAR-2003 (first entry)
DT
XX KIM-1 related protein, SEQ ID No 8.
DE
XX Cytostatic; gene therapy; antibody; antigen; antigen-binding;
KM proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
XX

KW renal disease; injury; renal cancer; human.
XX
XX Homo sapiens.
OS
XX WO200298920-A1.
PN
XX 12-DEC-2002.
PD
XX 31-MAY-2002; 2002WO-US017402.
XX
XX 01-JUN-2001; 2001US-0295449P.
PR 04-JUN-2001; 2001US-0295907P.
PR
XX (BIOJ) BIOGEN INC.
PA (GENO) GEN HOSPITAL CORP.
XX
XX Bailly V, Bonventre J;
PI
XX WPI; 2003-156845/15.
DR
XX
XX New antibody, antibody derivative or antigen-binding polypeptide that
PT inhibits proteolytic release of a soluble kidney injury molecule-1
PT polypeptide, useful for treating or preventing renal disease or injury,
PT e.g. renal cancer.
XX
XX Disclosure; Page 40-41; 42pp; English.
XX
XX The invention relates to a novel antibody, antibody derivative or antigen
CC -binding polypeptide that inhibits proteolytic release of a soluble
CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
CC The antibody, antibody derivative or antigen-binding polypeptide is
CC useful for treating or preventing renal disease or injury, e.g. renal
CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
CC protein of the invention
XX
XX Sequence 359 AA;
SQ
Query Match 100.0%; Score 99; DB 6; Length 359;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SSDGLMNNNOTQFLFHS 18
Db 264 SSDGLMNNNOTQFLFHS 281
RESULT 16
ADE36594
ID ADE36594 standard; protein; 359 AA.
XX
XX ADE36594;
AC
XX 29-JAN-2004 (first entry)
DT
XX Human NOV1b protein SEQ ID NO:4.
DE
XX human; NOVX; cytostatic; antiinflammatory; gene therapy; vaccine;
KM renal cancer; inflammation; tissue typing.
XX
XX Homo sapiens.
OS
XX WO2003080856-A2.
PN
XX 02-OCT-2003.
PD
XX 19-MAR-2003; 2003WO-US008490.
XX
XX 19-MAR-2002; 2002US-0365491P.
PR 13-SEP-2002; 2002US-0410618P.
PR
XX (CURA-) CURAGEN CORP.
PA
XX

PI Lepley DM, Rieger DK, Tse K, Rastelli L, Smithson G, Mesri M;
 PI Ooi CE, Anderson DM, Guo X, Glot L, Stalling G;
 XX WPI: 2003-876927/81.
 DR N-PSDB; ADE36593.
 XX
 PT New polypeptide, useful for preparing a composition for treating or
 PT preventing a pathology associated with NOVX polypeptide e.g. renal cancer
 PT or inflammation, or for tissue typing.
 XX
 PS Claim 1; SEQ ID NO 4; 239pp; English.
 XX
 CC The present invention describes an isolated human NOVX polypeptide, where
 CC X is 1a to 1d or 2a to 2h, or its mature form, a sequence that is at
 CC least 95 % identical to it, or a sequence comprising one or more
 CC conservative substitutions in the amino acid sequence. The NOVX
 CC polypeptide, and nucleic acid sequence encoding it, has cytosstatic and
 CC antiinflammatory activities, and can be used in gene therapy, and in
 CC vaccines. The NOVX polypeptide is useful for preparing a composition for
 CC treating or preventing a pathology associated with NOVX polypeptide e.g.
 CC renal cancer or inflammation, or for tissue typing. The present sequence
 CC represents human NOV1b from the present invention.
 CC
 XX Sequence 359 AA;
 SQ
 Query Match 100.0%; Score 99; DB 7; Length 359;
 Best Local Similarity 100.0%; Pred. No. 1,1e-06;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSDGLMNNNOTQFLFHS 18
 Db 264 SSDGLMNNNOTQFLFHS 281
 RESULT 17
 ADN38984
 ID ADN38984 standard; protein; 359 AA.
 XX
 AC ADN38984;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:302.
 XX
 KW Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cycostatic; cardiac; immunomodulatory;
 KW vulnerary; gene therapy; vaccine.
 KW
 XX Homo sapiens.
 OS
 PN MO2003042661-A2.
 XX
 PD 22-MAY-2003.
 XX
 DT 13-NOV-2002; 2002WO-US036810.
 XX
 PR 13-NOV-2001; 2001US-0350666P.
 PR 21-NOV-2001; 2001US-0322464P.
 PR 29-NOV-2001; 2001US-0334393P.
 PR 03-DEC-2001; 2001US-0335394P.
 PR 14-DEC-2001; 2001US-0340376P.
 PR 08-JAN-2002; 2002US-0347211P.
 PR 10-JAN-2002; 2002US-0355250P.
 PR 08-FEB-2002; 2002US-0356714P.
 PR 13-FEB-2002; 2002US-0359077P.
 PR 20-FEB-2002; 2002US-0368099P.
 PR 29-MAR-2002; 2002US-0370110P.
 PR 04-APR-2002; 2002US-0372246P.
 PR 12-APR-2002; 2002US-0372246P.

PR 05-JUN-2002; 2002US-0386614P.
 PR 16-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-0397775P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
 XX
 DR WPI: 2003-468649/44.
 DR N-PSDB; ADN38983.
 XX
 PT Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.
 XX
 PS Claim 12; SEQ ID NO 302; 1385pp; English.
 XX
 CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularisation syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a polypeptide of the invention.
 CC
 XX Sequence 359 AA;
 SQ
 Query Match 100.0%; Score 99; DB 7; Length 359;
 Best Local Similarity 100.0%; Pred. No. 1,1e-06;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSDGLMNNNOTQFLFHS 18
 Db 264 SSDGLMNNNOTQFLFHS 281
 RESULT 18
 ABP70443
 ID ABP70443 standard; protein; 364 AA.
 XX
 AC ABP70443;
 XX
 DT 22-APR-2003 (first entry)
 XX
 DE Amino acid sequence of human TIM-1 allele 6.
 XX
 KW T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;
 KW TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;
 KW myelodysplastic syndrome; airway hyperreactivity; cancer; asthma;
 KW allergic T cell response; autoimmune disease.
 KW
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT Misc-difference 123
 FT Misc-difference 161 /note= "encoded by GAG"
 FT Misc-difference 161 /note= "encoded by GACT"
 FT Misc-difference 201 /note= "encoded by ACAAGT"

XX WO2003002722-A2.
 PN 09-JAN-2003.
 XX
 PD 01-JUL-2002; 2002WO-US020890.
 XX
 PF 29-JUN-2001; 2001US-0302344P.
 XX
 PR (STRD) UNIV LELAND STANFORD JUNIOR.
 PA
 PI McIntire JJ, Dekryff RH, Umetsu DT, Freeman GJ, Kuchroo V;
 XX WPI; 2003-210268/20.
 DR N-PSDB; ABZ68336.
 XX
 PT New nucleic acid comprising a mammalian T cell immunoglobulin domain and
 PT mucin domain gene sequences, useful for treating cancer or asthma,
 PT allergy, eczema or autoimmune disease.
 XX
 PS Claim 10; Page 87-88; 94pp; English.
 XX
 CC The present sequence is a human T cell immunoglobulin domain and mucin
 CC domain-1 (TIM-1) polypeptide. The specification describes TIM-1, TIM-2,
 CC TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with
 CC conserved IGV and mucin domains. The locus comprising the TIM family is
 CC genetically associated with immune dysfunction, including asthma. The TIM
 CC gene family is located within a region of human chromosome 5 that is
 CC commonly deleted in malignancies and myelodysplastic syndrome. Variants
 CC of TIM-1 and TIM-3 are associated with susceptibility to airway
 CC hyperreactivity and allergic T cell responses, and other variants
 CC associated with protection against these responses. T cells express TIM
 CC proteins, which critically regulate CD4 T cell differentiation. Th1 cells
 CC preferentially express TIM-3, while Th2 cells preferentially express TIM-
 CC 1. TIM polypeptides and polynucleotides are useful for treating cancer,
 CC asthma, allergies, eczema or autoimmune diseases
 XX
 SQ Sequence 364 AA;
 XX
 QY
 DB 1 SSDGLMNNQTLFLHS 18
 269 SSDGLMNNQTLFLHS 286
 XX
 Query Match 100.0%; Score 99; DB 6; Length 364;
 Best Local Similarity 100.0%; Pred. No. 1.2e-06;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 19
 ABP70442
 ID ABP70442 standard; protein; 364 AA.
 AC ABP70442;
 XX
 DT 22-APR-2003 (first entry)
 XX
 DE Amino acid sequence of human TIM-1 allele 5.
 XX
 KM T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;
 KM TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;
 KM myelodysplastic syndrome; airway hyperreactivity; cancer; asthma;
 KM allergic T cell response; autoimmune disease.
 XX
 OS Homo sapiens.
 XX
 PN WO2003002722-A2.
 XX
 PD 09-JAN-2003.
 XX
 PF 01-JUL-2002; 2002WO-US020890.
 XX
 PR 29-JUN-2001; 2001US-0302344P.
 XX
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI McIntire JJ, Dekryff RH, Umetsu DT, Freeman GJ, Kuchroo V;
 XX WPI; 2003-210268/20.
 DR N-PSDB; ABZ68336.
 XX
 PT New nucleic acid comprising a mammalian T cell immunoglobulin domain and
 PT mucin domain gene sequences, useful for treating cancer or asthma,
 PT allergy, eczema or autoimmune disease.
 XX
 PS Claim 10; Page 86-87; 94pp; English.
 XX
 CC The present sequence is a human T cell immunoglobulin domain and mucin
 CC domain-1 (TIM-1) polypeptide. The specification describes TIM-1, TIM-2,
 CC TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with
 CC conserved IGV and mucin domains. The locus comprising the TIM family is
 CC genetically associated with immune dysfunction, including asthma. The TIM
 CC gene family is located within a region of human chromosome 5 that is
 CC commonly deleted in malignancies and myelodysplastic syndrome. Variants
 CC of TIM-1 and TIM-3 are associated with susceptibility to airway
 CC hyperreactivity and allergic T cell responses, and other variants
 CC associated with protection against these responses. T cells express TIM
 CC proteins, which critically regulate CD4 T cell differentiation. Th1 cells
 CC preferentially express TIM-3, while Th2 cells preferentially express TIM-
 CC 1. TIM polypeptides and polynucleotides are useful for treating cancer,
 CC asthma, allergies, eczema or autoimmune diseases
 XX
 SQ Sequence 364 AA;
 XX
 QY
 DB 1 SSDGLMNNQTLFLHS 18
 269 SSDGLMNNQTLFLHS 286
 XX
 Query Match 100.0%; Score 99; DB 6; Length 364;
 Best Local Similarity 100.0%; Pred. No. 1.2e-06;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

PA (STRD) UNIV LELAND STANFORD JUNIOR.
 XX
 PI McIntire JJ, Dekryff RH, Umetsu DT, Freeman GJ, Kuchroo V;
 XX WPI; 2003-210268/20.
 DR N-PSDB; ABZ68336.
 XX
 PT New nucleic acid comprising a mammalian T cell immunoglobulin domain and
 PT mucin domain gene sequences, useful for treating cancer or asthma,
 PT allergy, eczema or autoimmune disease.
 XX
 PS Claim 10; Page 86-87; 94pp; English.
 XX
 CC The present sequence is a human T cell immunoglobulin domain and mucin
 CC domain-1 (TIM-1) polypeptide. The specification describes TIM-1, TIM-2,
 CC TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with
 CC conserved IGV and mucin domains. The locus comprising the TIM family is
 CC genetically associated with immune dysfunction, including asthma. The TIM
 CC gene family is located within a region of human chromosome 5 that is
 CC commonly deleted in malignancies and myelodysplastic syndrome. Variants
 CC of TIM-1 and TIM-3 are associated with susceptibility to airway
 CC hyperreactivity and allergic T cell responses, and other variants
 CC associated with protection against these responses. T cells express TIM
 CC proteins, which critically regulate CD4 T cell differentiation. Th1 cells
 CC preferentially express TIM-3, while Th2 cells preferentially express TIM-
 CC 1. TIM polypeptides and polynucleotides are useful for treating cancer,
 CC asthma, allergies, eczema or autoimmune diseases
 XX
 SQ Sequence 364 AA;
 XX
 QY
 DB 1 SSDGLMNNQTLFLHS 18
 269 SSDGLMNNQTLFLHS 286
 XX
 Query Match 100.0%; Score 99; DB 6; Length 364;
 Best Local Similarity 100.0%; Pred. No. 1.2e-06;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 20
 ABP70440
 ID ABP70440 standard; protein; 365 AA.
 AC ABP70440;
 XX
 DT 22-APR-2003 (first entry)
 XX
 DE Amino acid sequence of human TIM-1 allele 3.
 XX
 KM T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;
 KM TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;
 KM myelodysplastic syndrome; airway hyperreactivity; cancer; asthma;
 KM allergic T cell response; autoimmune disease.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 208
 FT /note="encoded by RCG"
 PN WO2003002722-A2.
 XX
 PD 09-JAN-2003.
 XX
 PF 01-JUL-2002; 2002WO-US020890.
 XX
 PR 29-JUN-2001; 2001US-0302344P.
 XX
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI McIntire JJ, Dekryff RH, Umetsu DT, Freeman GJ, Kuchroo V;
 XX WPI; 2003-210268/20.
 DR

DR N-PSDB; ABZ68334.
 XX New nucleic acid comprising a mammalian T cell immunoglobulin domain and
 PT mucin domain gene sequences, useful for treating cancer or asthma,
 PT allergy, eczema or autoimmune disease.
 XX
 PS Claim 10; Page 83-84; 94pp; English.
 XX
 CC The present sequence is a human T cell immunoglobulin domain and mucin
 CC domain-1 (TIM-1) polypeptide. The specification describes TIM-1, TIM-2,
 CC TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with
 CC conserved IGV and mucin domains. The locus comprising the TIM family is
 CC genetically associated with immune dysfunction, including asthma. The TIM
 CC gene family is located within a region of human chromosome 5 that is
 CC commonly deleted in malignancies and myelodysplastic syndrome. Variants
 CC of TIM-1 and TIM-3 are associated with susceptibility to allergy
 CC hyperactivity and allergic T cell responses, and other variants
 CC associated with protection against these responses. T cells express TIM
 CC proteins, which critically regulate CD4 T cell differentiation. T11 cells
 CC preferentially express TIM-3, while Th2 cells preferentially express TIM-
 CC 1. TIM polypeptides and polynucleotides are useful for treating cancer,
 CC asthma, allergies, eczema or autoimmune diseases
 CC
 SQ Sequence 365 AA;
 Query Match 100.0%; Score 99; DB 6; Length 365;
 Best Local Similarity 100.0%; Pred. No. 1.2e-06;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 SSDGLMNNQOTLFLHS 18
 |||||
 Db 270 SSDGLMNNQOTLFLHS 287
 RESULT 21
 ABP70441
 ID ABP70441 standard; protein; 359 AA.
 XX
 AC ABP70441;
 XX
 DT 22-APR-2003 (first entry)
 XX
 DE Amino acid sequence of human TIM-1 allele 4.
 XX
 KW T cell immunoglobulin domain; mucin domain; TIM-1; TIM-2; TIM-3;
 KM TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;
 KM myelodysplastic syndrome; airway hyperactivity; cancer; asthma;
 KM allergic T cell response; autoimmune disease.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 172
 FT /note= "encoded by GACT"
 FT Misc-difference 202
 FT /note= "encoded by T"
 FT Misc-difference 240
 FT /note= "encoded by GCA"
 FT Misc-difference 270
 FT /note= "encoded by ART"
 XX
 PN WO2003002722-A2.
 XX
 PD 09-JAN-2003.
 XX
 PF 01-JUL-2002; 2002WO-US020890.
 XX
 PR 29-JUN-2001; 2001US-0302344P.
 XX
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 XX
 PI McIntire JJ, Dekruyff RH, Umetsu DT, Freeman GD, Kuchroo V;

DR WPI; 2003-210268/20.
 DR N-PSDB; ABZ68335.
 XX New nucleic acid comprising a mammalian T cell immunoglobulin domain and
 PT mucin domain gene sequences, useful for treating cancer or asthma,
 PT allergy, eczema or autoimmune disease.
 XX
 PS Claim 10; Page 84-85; 94pp; English.
 XX
 CC The present sequence is a human T cell immunoglobulin domain and mucin
 CC domain-1 (TIM-1) polypeptide. The specification describes TIM-1, TIM-2,
 CC TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with
 CC conserved IGV and mucin domains. The locus comprising the TIM family is
 CC genetically associated with immune dysfunction, including asthma. The TIM
 CC gene family is located within a region of human chromosome 5 that is
 CC commonly deleted in malignancies and myelodysplastic syndrome. Variants
 CC of TIM-1 and TIM-3 are associated with susceptibility to allergy
 CC hyperactivity and allergic T cell responses, and other variants
 CC associated with protection against these responses. T cells express TIM
 CC proteins, which critically regulate CD4 T cell differentiation. T11 cells
 CC preferentially express TIM-3, while Th2 cells preferentially express TIM-
 CC 1. TIM polypeptides and polynucleotides are useful for treating cancer,
 CC asthma, allergies, eczema or autoimmune diseases
 CC
 SQ Sequence 359 AA;
 Query Match 94.9%; Score 94; DB 6; Length 359;
 Best Local Similarity 94.4%; Pred. No. 6.6e-06;
 Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 SSDGLMNNQOTLFLHS 18
 |||||
 Db 264 SSDGLMNNQOTLFLHS 281
 RESULT 22
 AAR92803
 ID AAR92803 standard; protein; 451 AA.
 XX
 AC AAR92803;
 XX
 DT 24-MAY-1996 (first entry)
 XX
 DE Hepatitis A virus receptor.
 XX
 KW Hepatitis A virus; receptor; HAV; diagnosis; therapy; vaccine;
 KM transgenic animal.
 XX
 OS Cercopithecus aethiops.
 XX
 FN WO9604376-A1.
 FN
 PD 15-FEB-1996.
 XX
 PF 04-AUG-1995; 95WO-US009941.
 XX
 PR 05-AUG-1994; 94US-00287001.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Kaplan G, Feinstein SM,
 XX
 DR WPI; 1996-129392/13.
 DR N-PSDB; AAT17864.
 XX
 PT DNA encoding Hepatitis A virus receptor and its fragments - useful for
 PT the detection, diagnosis, prevention and treatment of HAV infection.
 XX
 PS Claim 1; Page 39-41; 52pp; English.
 XX
 CC A hepatitis A virus (HAV) receptor (AAR92803) was identified as the
 CC product of a nucleic acid sequence isolated from HAV-infected primary
 CC African green monkey kidney cells. Recombinant receptor can be produced

CC in large amounts by expression of the nucleic acid in prokaryotic or
 CC eukaryotic host cells or in transgenic animals. It can be used in the
 CC detection, prevention and treatment of HAV infection, in methods of
 CC purifying or removing HAV from samples, and methods of determining the
 CC anti-HAV activity of cpds. Expression in transgenic animals provides a
 CC method for testing vaccine efficacy

XX Sequence 451 AA;

Query Match 84.8%; Score 84; DB 2; Length 451;
 Best Local Similarity 88.9%; Pred. No. 0.00028;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SSDGLMNNQTLFLEHS 18
 |||||
 Db 344 SSDGLMNNQTLSPFHS 361

RESULT 23

AAO26686 standard; peptide; 18 AA.

XX AAO26686;

DT 20-MAR-2003 (first entry)

XX Human KIM-1 mucin domain related peptide, SEQ ID No 14.

XX Cycostatic; gene therapy; antibody; antigen; antigen-binding;
 KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
 KW renal disease; injury; renal cancer.

XX Synthetic.

XX WO200298920-A1.

XX 12-DEC-2002.

XX 31-MAY-2002; 2002WO-US017402.

XX 01-JUN-2001; 2001US-0295449P.

PR 04-JUN-2001; 2001US-0295907P.

XX (BIOJ) BIOGEN INC.
 PA (GEHO) GEN HOSPITAL CORP.

PI Bailly V, Bonventure J;

XX WPI; 2003-156845/15.

XX New antibody, antibody derivative or antigen-binding polypeptide that
 PT inhibits proteolytic release of a soluble kidney injury molecule-1
 PT polypeptide, useful for treating or preventing renal disease or injury,
 PT e.g. renal cancer.

XX Disclosure; Fig 1A; 42pp; English.

XX The invention relates to a novel antibody, antibody derivative or antigen
 CC binding polypeptide that inhibits proteolytic release of a soluble
 CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
 CC The antibody, antibody derivative or antigen-binding polypeptide is
 CC useful for treating or preventing renal disease or injury, e.g. renal
 CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
 CC polypeptide. This sequence represents a human KIM-1 mucin domain related
 CC peptide of the invention

XX Sequence 18 AA;

Query Match 53.5%; Score 53; DB 6; Length 18;
 Best Local Similarity 100.0%; Pred. No. 0.45;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSDGLMNNN 9

Db |||||
 10 SSDGLMNNN 18

RESULT 24

AAE18306 standard; protein; 981 AA.

XX AAE18306;

DT 07-MAY-2002 (first entry)

XX Venezuelan equine encephalitis virus (VEE) glycoprotein V3014.

XX Venezuelan equine encephalitis virus; VEE; vaccine; HIV infection;
 KW human immunodeficiency virus; alpha-virus replicon; immune response;
 KW therapy; glycoprotein V3014.

XX Venezuelan equine encephalitis virus.

XX WO200203917-A2.

XX 17-JAN-2002.

XX 09-JUL-2001; 2001WO-US021701.

XX 07-JUL-2000; 2000US-0216995P.

XX (ALPH-) ALPHA-VAX INC.
 PA (UTNC-) UNIV NORTH CAROLINA.

PI Olmsted R, Keith P, Dryga S, Caley I, Maughan M, Johnston R;
 PI Davis N, Swanson R;

XX WPI; 2002-171664/22.

DR N-PSDB; AAD29138.

XX Composition useful for treating or preventing HIV infections, comprises
 PT two or more isolated nucleic acids encoding env, gag or pol gene product
 PT of HIV or immunogenic fragment of the gene products.

XX Example 4; Page 180-183; 20pp; English.

XX The invention relates to a composition comprising isolated nucleic acids
 CC encoding env, gag or pol gene product of human immunodeficiency virus or
 CC immunogenic fragment of the gene products. The gag gene product is
 CC modified to inhibit formation of virus-like particles containing gag gene
 CC product and their release from cells, and the pol gene product is
 CC modified to inhibit reverse transcriptase activity. The invention also
 CC relates to a method for producing an alpha-virus replicon particle used
 CC in vaccines. The composition is useful for inducing an immune response to
 CC human immunodeficiency virus (HIV) or for treating or preventing HIV
 CC infection in a subject. The alpha-virus replicon particle is useful in a
 CC vaccine. The composition is useful for administering a protein or peptide
 CC to a subject. A composition comprising heparin affinity-purified alpha-
 CC virus replicon particle is useful as a clinical trial material and as a
 CC commercial product. The present sequence is Venezuelan equine
 CC encephalitis virus (VEE) glycoprotein V3014 used in the exemplification
 CC of the invention

XX Sequence 981 AA;

Query Match 48.5%; Score 48; DB 5; Length 981;
 Best Local Similarity 50.0%; Pred. No. 1.8e+02;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SSDGLMNNQTLFLE 16
 |||||
 Db 489 SLDFLMNNQMFMIQ 504

RESULT 25

AAO26688

ID AAO26688 standard; peptide; 18 AA.
XX
AC AAO26688;
XX
DT 20-MAR-2003 (first entry)
XX
DE Human KIM-1 mucin domain related peptide, SEQ ID No 16.
XX
KW Cytostatic; gene therapy; antibody; antigen; antigen-binding;
KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
KW renal disease; injury; renal cancer.
XX
OS Synthetic.
XX
PN WO200298920-A1.
XX
PD 12-DEC-2002.
XX
PF 31-MAY-2002; 2002WO-US017402.
XX
PR 01-JUN-2001; 2001US-0285449P.
PR 04-JUN-2001; 2001US-0293907P.
XX
PA (BIOJ) BIOGEN INC.
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Bailly V, Bonventre J;
XX
DR WPI; 2003-156845/15.
XX
PT New antibody, antibody derivative or antigen-binding polypeptide that
PT inhibits proteolytic release of a soluble kidney injury molecule-1
PT polypeptide, useful for treating or preventing renal disease or injury,
PT e.g. renal cancer.
XX
PS Disclosure; Fig 1A; 42pp; English.
XX
CC The invention relates to a novel antibody, antibody derivative or antigen
CC binding polypeptide that inhibits proteolytic release of a soluble
CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
CC The antibody, antibody derivative or antigen-binding polypeptide is
CC useful for treating or preventing renal disease or injury, e.g. renal
CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
CC polypeptide. This sequence represents a human KIM-1 mucin domain related
CC peptide of the invention
XX
SQ Sequence 18 AA;
XX
Query Match 46.5%; Score 46; DB 6; Length 18;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 QTOFLFHS 18
DB 1 QTOFLFHS 9
XX
RESULT 26
ABG93049
ID ABG93049 standard; protein; 508 AA.
XX
AC ABG93049;
XX
DT 21-NOV-2002 (first entry)
XX
DE S. cerevisiae BAX-associated protein fragment SEQ ID 56.
XX
KW Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;
KW vacuotropic; vaccine; gene therapy; proliferative disorder; cancer;
KW apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;
KW neurodegeneration; cell death.
XX
OS Saccharomyces cerevisiae.

XX
PN WO200264766-A2.
XX
PD 22-AUG-2002.
XX
PF 21-DEC-2001; 2001WO-EP015398.
XX
PR 22-DEC-2000; 2000EP-00870318.
PR 04-JAN-2001; 2001EP-00870002.
PR 09-JAN-2001; 2001EP-00870003.
XX
PA (JANC) JANSSEN PHARM NV.
XX
PI Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;
XX
DR WPI; 2002-667002/71.
DR N-PSDB; ABQ76315.
XX
PT New isolated nucleic acid representing a synthetic BAX-gene, useful as
PT medication for treating, preventing and/or alleviating yeast or fungal
PT infections or proliferative disorders, or for preventing apoptosis in
PT certain diseases.
XX
PS Claim 36; Fig 1; 344pp; English.
XX
CC This invention describes a novel nucleic acid representing a synthetic
CC Bax gene. The Bax gene of the invention is useful for identifying Bax-
CC resistant yeast or fungi, identifying, or obtaining and identifying
CC Candida spp. sequences that are differentially expressed in a pathway
CC eventually leading to programmed cell death or identifying inhibitors or
CC inhibitor sequences of Bax-induced cell death. The products of the
CC invention have cytostatic, fungicide, immunosuppressive, virucide and
CC vacuotropic activity and can be used in vaccines or for gene therapy. The
CC isolated nucleic acids, polypeptides, pharmaceutical compositions,
CC antisense molecules and antibodies are useful as medicaments or in
CC preparing a medicament for treating, preventing and/or alleviating
CC diseases associated with yeast or fungi or proliferative disorders, such
CC as cancer, or for preventing apoptosis in certain diseases. The compounds
CC or polypeptides, or the genetically modified organism are useful for
CC preparing a medicament for modifying the endogenous flora of humans and
CC other mammals. The vaccine is useful for immunising against yeast or
CC fungal infections. Apoptosis-related diseases include autoimmune disease,
CC ischaemia, diseases related with viral infections or neurodegenerations.
CC This sequence represents a polypeptide associated with the Bax gene
CC described in the disclosure of the invention
XX
SQ Sequence 508 AA;
XX
Query Match 46.5%; Score 46; DB 5; Length 508;
Best Local Similarity 63.6%; Pred. No. 1.8e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 6 WNNNQTOFLFLE 16
DB 160 WNNNASWIMFLE 170
XX
RESULT 27
ABR52836
ID ABR52836 standard; protein; 508 AA.
XX
AC ABR52836;
XX
DT 20-JUN-2003 (first entry)
XX
DE Protein sequence #SEQ ID 537.
XX
KW Multiprotein complex; eukaryote; drug target; diagnosis.
KW Saccharomyces cerevisiae.
XX
PN EPI258494-A1.
XX

PD 20-NOV-2002.
XX
XX 20-DEC-2001; 2001EP-00130253.
XX
PR 15-MAY-2001; 2001EP-00111774.
XX
XX (CELL-) CELTZOME AG.
XX
PI Bauer A, Gavin A, Grandi P, Krause R, Kruse UD, Kuester BD;
PI Marzlioch M, Schultze JD, Superti-Furga GD;
XX
XX WPI; 2003-250078/25.
DR N-PSDB; ACC60878.
XX
XX New isolated protein complexes useful for diagnosing a disease or
PT disorder, or as a target for an active agent of a pharmaceutical,
PT preferably a drug target in the treatment or prevention of disease or
PT disorder.
XX
XX Disclosure; SEQ ID NO 537; 17bp + Sequence listing; English.
XX
CC The invention relates to multiprotein complexes from eukaryotes. Proteins
CC of the invention and DNA sequences encoding them are given in records
CC AB52568-ABR33903 and ACC60610-ACC61944 respectively. The complexes are
CC obtainable by using a protein as a bait and isolating the set of proteins
CC which is attached thereto from cells. Such protein complexes may comprise
CC up to 30 distinct proteins. Protein complexes of the invention are useful
CC for diagnosing a disease or disorder, or as a target for an active agent
CC of a pharmaceutical, preferably a drug target in the treatment or
CC prevention of a disease or disorder. Note: The sequence data for this
CC patent is not represented in the printed specification, but is based on
CC sequence information supplied by the European Patent Office. The complete
CC document is available on CD-ROM
XX
SQ Sequence 508 AA;

Query Match 46.5%; Score 46; DB 6; Length 508;
Best Local Similarity 63.6%; Pred. No. 1.8e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 WNNNGTQLE 16
Db 160 WNNNASWIFLE 170

RESULT 28
ADK62270
ID ADK62270 standard; protein; 508 AA.
XX
AC ADK62270;
XX
DT 06-MAY-2004 (first entry)
XX
DE Disease treating protein complex-derived protein #261.
XX
XX protein complex; drug target; diagnosis.
XX
XX Unidentified.
XX
XX EP1338608-A2.
XX
XX 27-AUG-2003.
XX
XX 20-DEC-2002; 2002EP-00102902.
XX
XX 20-DEC-2001; 2001EP-00130253.
XX
XX (CELL-) CELTZOME AG.
XX
XX Bauer A, Gavin A, Superti-Furga G, Kuester B, Schultze J;
XX Marzlioch M, Grandi P, Krause R, Kruse U, Merino A, Bauch A;
XX Michon A, Leutwein C, Rick J;
XX

DR WPI; 2003-638460/61.
DR N-PSDB; ADK62271.
XX
XX New screening and protein complexes from eukaryotes, useful as targets in
PT drug screening, or in diagnosing or screening for the presence of a
PT disease or disorder, or a predisposition for developing a disease or
PT disorder in a subject.
XX
XX Disclosure; SEQ ID NO 521; 13bp; English.
XX
CC The invention relates to novel protein complexes comprising a first and a
CC second protein, or its derivative, fragment, homologue or variant. The
CC proteins are selected from given protein complexes, which are not defined
CC in the specification. The variants are encoded by nucleic acids that
CC hybridize to the nucleic acids encoding the proteins under low stringency
CC conditions. The protein complexes are useful as targets for an active
CC agent of a pharmaceutical. These protein complexes are particularly
CC useful as drug targets for the treatment or preventing of a disease or
CC disorder. The complexes and methods above are useful in diagnosing or
CC screening for the presence of a disease or disorder or a predisposition
CC for developing a disease or disorder in a subject. These are also useful
CC in screening for a drug for treatment or prevention of a disease or
CC disorder. The molecule that modulates the amount, activity or protein
CC components of the complex is useful for the manufacture of a medicament
CC for the treatment or prevention of a disease or disorder. This sequence
CC corresponds to a protein of the invention. (Note: the sequence data for
CC this patent did not form part of the printed specification but was
XX obtained from the EPO in electronic format).
XX
SQ Sequence 508 AA;

Query Match 46.5%; Score 46; DB 7; Length 508;
Best Local Similarity 63.6%; Pred. No. 1.8e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 WNNNGTQLE 16
Db 160 WNNNASWIFLE 170

RESULT 29
ABG71366
ID ABG71366 standard; protein; 545 AA.
XX
XX ABG71366;
XX
XX 28-JAN-2003 (first entry)
XX
XX Potato plastidic targeted beta-amylase (ppt-beta-amylase).
XX
XX Potato; starch degrading enzyme; enzyme; starch degradation;
XX washing agent; flushing agent; potato plastidic targeted beta-amylase;
XX ppt-beta-amylase.
XX
XX Solanum tuberosum.
XX
XX WO200286112-A2.
XX
XX 31-OCT-2002.
XX
XX 25-APR-2002; 2002WO-EP004609.
XX
XX 25-APR-2001; 2001EP-00110005.
XX
XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX
XX Scheidig A, Koesmann J, Froehlich A;
XX
XX WPI; 2003-058726/05.
XX
XX N-PSDB; ABS56813.
XX
XX Novel nucleic acid useful for modifying starch degradation and as
PT hybridization probes to isolate related genes, encodes enzyme involved in

PT starch degradation.
 XX
 PS Claim 1, Page 105-107; 107pp; English.
 CC The invention relates to a nucleic acid molecule encoding a starch
 CC degrading enzyme. The nucleic acid is useful for reducing endogenous
 CC activity of a biologically active fragment, for modifying starch
 CC degradation, as a hybridisation probe to isolate related genes and as a
 CC primer for amplification techniques such as polymerase chain reaction.
 CC The nucleic acid is also useful as an agent for starch degradation in a
 CC washing or flushing agent. This sequence represents a potato plastidic
 CC targeted beta-amylase (ppp-beta-amylase) polypeptide
 XX
 SQ Sequence 545 AA;
 Query Match 46.5%; Score 46; DB 6; Length 545;
 Best Local Similarity 57.1%; Pred. No. 2e+02;
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 3 DGLMNNNOTLPLE 16
 324 DGTWNSYGFPLE 337
 RESULT 30
 ID ADJ49214 standard; protein; 545 AA.
 AC ADJ49214;
 DT 06-MAY-2004 (first entry)
 XX
 DE Oil-associated gene related protein #714.
 XX
 KM oil-associated gene; transgenic; enhanced seed oil; vegetable oil.
 XX
 OS Unidentified.
 XX
 PN US2004025202-A1.
 XX
 PD 05-FEB-2004.
 XX
 PF 14-MAR-2003; 2003US-00389566.
 XX
 PR 15-MAR-2002; 2002US-0365301P.
 PR 26-JUN-2002; 2002US-0391786P.
 PR 26-JUN-2002; 2002US-0392018P.
 PA (LAURIE C C.
 PA (RAVANELLO M.
 PA (SAVAGE T.
 PA (LEDEAUX J R.
 PA (ROGERS J A.
 PI Laurie CC, Ravello M, Savage T, Ledeaux JR, Rogers JA;
 PI WPI; 2004-142683/14.
 DR
 XX
 PT Novel recombinant DNA construct comprising a promoter functional in
 PT plants operably linked to an oil-associated gene for producing transgenic
 PT plant seed.
 XX
 XX Example 3; SEQ ID NO 1218; 22pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in plants operably linked to an oil-associated gene.
 CC The construct is useful for transgenic plant seed which has in its genome
 CC the construct, that is functional in the plant to transcribe the oil-
 CC associated gene. The transgenic plant seed grows into a plant having
 CC enhanced seed oil as compared to wild type. The construct is useful for
 CC producing hybrid maize seed. The transgenic plant seed is useful for
 CC producing vegetable oil. The present sequence represents the amino acid
 CC sequence of an oil-associated gene related protein.

XX
 SQ Sequence 545 AA;
 Query Match 46.5%; Score 46; DB 8; Length 545;
 Best Local Similarity 57.1%; Pred. No. 2e+02;
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 3 DGLMNNNOTLPLE 16
 324 DGTWNSYGFPLE 337
 RESULT 31
 ID ABR57094 standard; peptide; 114 AA.
 AC ABR57094;
 DT 26-AUG-2003 (first entry)
 XX
 DE MLHR comparison related RAHEPLEC2 amino acid sequence.
 XX
 KM LHR; MLHR; immunoglobulin, Ig; immunoglobulin heavy chain dimer;
 KM ligand-combined partner; cell surface adhesion molecule; ligand;
 KM lymphocyte cell surface glycoprotein.
 XX
 OS Unidentified.
 XX
 PN JP2002325589-A.
 XX
 PD 12-NOV-2002.
 XX
 PF 21-NOV-1990; 2002JP-00026825.
 XX
 PR 22-NOV-1989; 89US-00444625.
 PR 21-NOV-1990; 91JP-00501520.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PD WPI; 2003-407283/39.
 DR
 XX
 PT A fused protein consisting of a ligand-combined protein and a stable
 PT plasma protein, a polypeptide, a nucleic acid, a replicable expression
 PT vector, a composition.
 XX
 XX Example 1, Fig 11; 44pp; Japanese.
 PS
 CC The present invention describes a polypeptide (I) containing an
 CC immunoglobulin (Ig) heavy chain dimer having no Ig light chain in which
 CC the amino acid sequence of a ligand-combined partner is a receptor, a
 CC carrier protein, a hormone, a growth factor, an enzyme or a nutritive
 CC substance, but is not a subunit polypeptide encoded by a lymphocyte
 CC inducing receptor, an Ig gene super family-constituting element, a
 CC protein homologous to it, or a separated gene substitute of the variable
 CC region of at least one Ig heavy chain. Its combined partner is fused with
 CC the amino acid sequence of an Ig stationary region at its C-end and
 CC maintains its combining feature. Also described: (1) a nucleic acid
 CC encoding (I); (2) a replicable expression vector containing the nucleic
 CC acid of (1); (3) a composition containing cells transformed by the
 CC expression vector of (2); and (4) producing (1), comprising culturing the
 CC cells of (3) in which the transformed cells are cultured and the
 CC polypeptide is recovered from the cell culture. The polypeptide is useful
 CC as a cell surface adhesion molecule and a ligand. In an example from the
 CC present invention mouse lymphocyte cell surface glycoprotein designated
 CC LHR (MLHR) was isolated from a mouse spleen and purified and cloned. The
 CC present sequence represents an amino acid sequence given in comparison
 CC with an MLHR amino acid sequence, which is used in an example from the
 CC present invention
 XX
 SQ Sequence 114 AA;
 Query Match 45.5%; Score 45; DB 6; Length 114;
 Best Local Similarity 72.7%; Pred. No. 53;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 SDGLMNNQTO 12
|||...|
Db 97 SDGLMNDNFCQ 107

RESULT 32
ABR57095
ID ABR57095 standard; peptide; 114 AA.
XX
AC ABR57095;
XX
DT 26-AUG-2003 (first entry)
XX
DE MLHR comparison related RA.ASGREC amino acid sequence.
XX
KW MLHR; MLHR; immunoglobulin; Ig; immunoglobulin heavy chain dimer;
KW ligand-combined partner; cell surface adhesion molecule; ligand;
KW lymphocyte cell surface glycoprotein.
XX
OS Unidentified.
XX
PN JP2002325589-A.
XX
PD 12-NOV-2002.
XX
PF 21-NOV-1990; 2002JP-00026825.
XX
PR 22-NOV-1989; 89US-00444625.
PR 21-NOV-1990; 91JP-00501520.
XX
XX
PA (GETH) GENENTECH INC.
XX
DR WPI; 2003-407283/39.
XX
PT A fused protein consisting of a ligand-combined protein and a stable
PT plasma protein, a polypeptide, a nucleic acid, a replicable expression
PT vector, a composition.
XX
PS Example 1; Fig 11; 44pp; Japanese.
XX
XX The present invention describes a polypeptide (I) containing an
CC immunoglobulin (Ig) heavy chain dimer having no Ig light chain in which
CC the amino acid sequence of a ligand-combined partner is a receptor, a
CC carrier protein, a hormone, a growth factor, an enzyme or a nutritive
CC substance, but is not a subunit polypeptide encoded by a lymphocyte
CC inducing receptor, an Ig gene super family-constituting element, a
CC protein homologous to it, or a separated gene substitute of the variable
CC region of at least one Ig heavy chain. Its combined partner is fused with
CC the amino acid sequence of an Ig stationary region at its C-end and
CC maintains its combining feature. Also described: (1) a nucleic acid
CC encoding (I); (2) a replicable expression vector containing the nucleic
CC acid of (1); (3) a composition containing cells transformed by the
CC expression vector of (2); and (4) producing (1), comprising culturing the
CC cells of (3) in which the transformed cells are cultured and the
CC polypeptide is recovered from the cell culture. The polypeptide is useful
CC as a cell surface adhesion molecule and a ligand. In an example from the
CC present invention mouse lymphocyte cell surface glycoprotein designated
CC LHR (MLHR) was isolated from a mouse spleen and purified and cloned. The
CC present sequence represents an amino acid sequence given in comparison
CC with an MLHR amino acid sequence, which is used in an example from the
CC present invention
XX
SQ Sequence 114 AA;

Query Match 45.5%; Score 45; DB 6; Length 114;
Best Local Similarity 72.7%; Pred. No. 53;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 SDGLMNNQTO 12
|||...|
Db 97 SDGLMNDNFCQ 107

RESULT 33
ID AAM42107
XX AAM42107 standard; protein; 244 AA.
XX
AC AAM42107;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 7038.
XX
XX Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
XX (HYSE-) HYSEQ INC.
XX
PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou F, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
DR N-PSDB; AAI61263.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
PT
XX
XX Example 2; SEQ ID NO 7038; 10078pp; English.
PS
XX The invention relates to human nucleic acids (AA157798-AA161369) and the
CC encoded polypeptides (AAM38642-AAM4213) with noctropic;
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 244 AA;

Query Match 45.5%; Score 45; DB 4; Length 244;
Best Local Similarity 47.4%; Pred. No. 1.2e+02;
Matches 9; Conservative 3; Mismatches 3; Indels 4; Gaps 1;

RESULT 36
 ID ADD46355 standard; protein; 301 AA.
 XX
 AC ADD46355;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Rat Protein P08290, SEQ ID NO 12033.
 XX
 KM Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
 XX
 OS Rattus norvegicus.
 XX
 PN W02003016475-A2.
 XX
 PD 27-FEB-2003.
 XX
 PF 14-AUG-2002; 2002WO-US025765.
 XX
 PR 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX
 PA (GENO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX
 PI Woolf C, D'urso D, Befort K, Costigan M;
 XX
 DR WPI; 2003-268312/26.
 DR GENBANK; P08290.
 XX
 PT New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX
 PS Claim 1; Page; 1017P; English.
 XX
 CC The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 301 AA;

Query Match 45.5%; Score 45; DB 7; Length 301;
 Best Local Similarity 72.7%; Pred. No. 1.5e+02;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SDGLMNNQTO 12
 Db 276 SDGLMNDNFCQ 286
 RESULT 37
 ID ADE63357 standard; protein; 301 AA.
 XX
 AC ADE63357;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Rat Protein P08290, SEQ ID NO 9295.
 XX
 KM Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
 XX
 OS Rattus norvegicus.
 XX
 PN W02003016475-A2.
 XX
 PD 27-FEB-2003.
 XX
 PF 14-AUG-2002; 2002WO-US025765.
 XX
 PR 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX
 PA (GENO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX
 PI Woolf C, D'urso D, Befort K, Costigan M;
 XX
 DR WPI; 2003-268312/26.
 DR GENBANK; P08290.
 XX
 PT New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX
 PS Claim 1; Page; 1017P; English.
 XX
 CC The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 301 AA;

Query Match 45.5%; Score 45; DB 7; Length 301;
Best Local Similarity 72.7%; Pred. No. 1.5e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 SDGLMNNNQ 12
|||
Db 276 SDGLMNDNFCQ 266

RESULT 38

AAB92470
ID AAB92470 standard; protein; 309 AA.

XX AAB92470;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:10536.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

XX EPI074617-A2.

XX 07-FEB-2001.

PD 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.

PR 17-AUG-1999; 99JP-00300253.

PR 11-JAN-2000; 2000JP-00118776.

PR 02-MAY-2000; 2000JP-00183767.

PR 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

PI Ota T, Iwagai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

PS Claim 8; SEQ ID NO 10536; 2537bp + Sequence Listing; English.

CC The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises:

CC (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602

CC nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the

CC specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; AAH13629 to AAH13632 represent

CC oligonucleotides, all of which are used in the exemplification of the present invention

XX Sequence 309 AA;

Query Match 45.5%; Score 45; DB 4; Length 309;
Best Local Similarity 47.4%; Pred. No. 1.5e+02;
Matches 9; Conservative 3; Mismatches 3; Indels 4; Gaps 1;

QY 3 DGLMNNNQ---TQLFLEH 17
|||
Db 74 DGTWNNNQLEMAQURKH 92

RESULT 39

AAB97591
ID AAB97591 standard; protein; 309 AA.

XX AAB97591;

DT 27-JUN-2002 (first entry)

DE Novel human protein SEQ ID NO: 859.

XX Human; anti-anaemic; vulnerary; anti-inflammatory; immunomodulator;

XX anti-infectivity; cerebroprotective; cytostatic; rheumatic; gene therapy;

XX neuroprotective; antiparkinsonian; protein therapy; EST;

XX expressed sequence tag.

OS Homo sapiens.

XX WO200222660-A2.

XX 21-MAR-2002.

XX 10-SEP-2001; 2001WO-US026015.

XX 11-SEP-2000; 2000US-00659671.

XX (HYSE-) HYSEQ INC.

PI Tang YF, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

PI Xue AJ, Yang Y, Wehrman T, Drmanac R;

XX WPI; 2002-292408/33.

XX N-PSDB; ABB32777.

PS Claim 20; SEQ ID NO 859; 509bp; English.

CC The present invention provides the protein and coding sequences of 444 novel human proteins. These were isolated from expressed sequences tags

CC (ESTs). They can be used to stimulate cell growth, to regulate haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth

CC e.g. in burn treatment, to regulate the immune system e.g. to treat multiple sclerosis, to regulate activin or inhibin e.g. to treat

CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke and cancer, to screen for drugs, to treat inflammatory conditions e.g.

CC rheumatoid arthritis, and to treat nervous system disorders e.g. Parkinson's disease. The present sequence is a protein of the invention

XX Sequence 309 AA;

Query Match 45.5%; Score 45; DB 5; Length 309;
Best Local Similarity 47.4%; Pred. No. 1.5e+02;
Matches 9; Conservative 3; Mismatches 3; Indels 4; Gaps 1;

QY 3 DGLMNNNQ---TQLFLEH 17
|||
Db 74 DGTWNNNQLEMAQURKH 92

RESULT 40

AAM40322
ID AAM40322 standard; protein; 511 AA.

Search completed: December 30, 2004, 16:10:01
Job time : 165 secs

```

XX  AAM40322;
XX
XX  22-OCT-2001 (first entry)
XX
XX  Human polypeptide SEQ ID NO 3467.
XX
XX
XX  Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX  peripheral nervous system; neuropathy; central nervous system; CNS;
XX  Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX  amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX  chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX  leukaemia.
XX
XX  Homo sapiens.
XX
XX  WO200153312-A1.
XX
XX  26-JUL-2001.
XX
XX  26-DEC-2000; 2000WO-US034263.
XX
XX  23-DEC-1999; 99US-00471275.
XX  21-JAN-2000; 2000US-00488725.
XX  25-APR-2000; 2000US-00552317.
XX  20-JUN-2000; 2000US-00598042.
XX  19-JUL-2000; 2000US-00620312.
XX  03-AUG-2000; 2000US-00653450.
XX  14-SEP-2000; 2000US-00662191.
XX  19-OCT-2000; 2000US-00693036.
XX  29-NOV-2000; 2000US-00727344.
XX
XX  (HYSE-) HYSEQ INC.
XX
XX  Tang YT, Liu C, Aundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX  Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
XX  Zhou F, Goodrich R, Drmanac RT;
XX  N-PSDB; AAI59476.
XX
XX  WPI; 2001-442253/47.
XX
XX  Novel nucleic acids and polypeptides, useful for treating disorders such
XX  as central nervous system injuries.
XX
XX  Example 6; SEQ ID NO 3467; 10078pp; English.
XX
XX  The invention relates to human nucleic acids (AA157798-AA161369) and the
XX  encoded polypeptides (AAM3642-AAM42213) with nootropic,
XX  immunosuppressant and cytostatic activity. The polynucleotides are useful
XX  in gene therapy. A composition containing a polypeptide or polynucleotide
XX  of the invention may be used to treat diseases of the peripheral nervous
XX  system, such as peripheral nervous injuries, peripheral neuropathy and
XX  localised neuropathies and central nervous system diseases, such as
XX  Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX  lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX  utilisation of the activities such as: Immune system suppression,
XX  Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX  and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX  assays for receptor activity, arthritis and inflammation, leukemias and
XX  C.N.S disorders. Note: The sequence data for this patent did not form
XX  part of the printed specification
XX
XX  Sequence 511 AA;
XX
Query Match          45.5%; Score 45; DB 4; Length 511;
Best Local Similarity 47.4%; Pred. NO. 2.6e+02;
Matches 9; Conservative 3; Mismatches 3; Indels 4; Gaps 1;
QY  3 DGLMNNNQ---TQLEH 17
    |||||:|:|
Db  74 DGTWNDNQLEMAQRLRIKH 92

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 16:13:23 : Search time 143 Seconds
(without alignments)
45.280 Million cell updates/sec

Title: US-10-718-321-1

Perfect score: 99

Sequence: 1 SSDGLMNNQTLFLHS 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1599051 seqs, 359727711 residues

Total number of hits satisfying chosen parameters: 1599051

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubppaa/US07_PUBSCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBSCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/PCITUS_PUBSCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBSCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US08A_PUBSCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBSCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBSCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US10_PUBSCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US10A_PUBSCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBSCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppaa/US10D_PUBSCOMB.pep:*
17: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubppaa/US60_PUBSCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	99	100.0	359 14 US-10-188-012-17	Sequence 17, Appl
2	99	100.0	359 14 US-10-188-012-19	Sequence 19, Appl
3	99	100.0	359 14 US-10-235-027-102	Sequence 302, App
4	99	100.0	359 15 US-10-188-832-64	Sequence 64, Appl
5	99	100.0	364 14 US-10-188-012-25	Sequence 25, Appl
6	99	100.0	364 14 US-10-188-012-27	Sequence 27, Appl
7	99	100.0	365 14 US-10-188-012-21	Sequence 21, Appl
8	94	94.9	359 14 US-10-188-012-23	Sequence 23, Appl
9	48	48.5	981 9 US-09-991-258-13	Sequence 13, Appl
10	47	47.5	102 15 US-10-424-599-175126	Sequence 175126, Sequence 1469, Ap
11	46	46.5	508 14 US-10-369-493-1469	Sequence 56, Appl
12	46	46.5	508 16 US-10-451-467A-56	Sequence 56, Appl
13	46	46.5	545 15 US-10-389-566-1218	Sequence 1218, Ap

14	45.5	384	16 US-10-437-963-163657	Sequence 163657,
15	45.5	770	16 US-10-437-963-163658	Sequence 163658,
16	45.5	273	14 US-10-156-761-9511	Sequence 9511, Ap
17	45.5	687	15 US-10-282-122A-48522	Sequence 48522, A
18	45.5	1464	10 US-09-934-070-15	Sequence 15, Appl
19	45.5	1464	10 US-09-922-011-1	Sequence 1, Appl1
20	45.5	1464	10 US-10-222-772-15	Sequence 15, Appl
21	44.4	243	17 US-10-739-930-9140	Sequence 9140, Ap
22	44.4	268	15 US-10-424-599-173377	Sequence 173377,
23	44.4	432	15 US-10-282-122A-57369	Sequence 57369, A
24	44.4	539	14 US-10-369-493-3911	Sequence 3911, Ap
25	44.4	590	10 US-09-826-509-517	Sequence 517, App
26	44.4	590	10 US-10-225-567A-440	Sequence 440, Appl
27	44.4	590	14 US-10-352-684A-40	Sequence 11, Appl
28	44.4	608	13 US-10-029-009-11	Sequence 23, Appl
29	44.4	626	13 US-10-029-009-23	Sequence 2361, A
30	44.4	1042	14 US-10-369-493-23561	Sequence 282496,
31	43.4	48	15 US-10-424-599-282496	Sequence 174021,
32	43.4	100	15 US-10-424-599-174021	Sequence 312882,
33	43.4	116	17 US-10-425-115-312882	Sequence 65, Appl
34	43.4	233	15 US-10-321-925-65	Sequence 176868,
35	43.4	554	16 US-10-437-963-176868	Sequence 8, Appl1
36	43.4	561	14 US-10-094-468-8	Sequence 174682,
37	43.4	684	16 US-10-437-963-174682	Sequence 9394, Ap
38	43.4	973	14 US-10-156-761-9394	Sequence 265756,
39	42.4	42	15 US-10-424-599-265756	Sequence 367358,
40	42.4	50	17 US-10-425-115-367358	Sequence 364078,
41	42.4	151	17 US-10-425-115-364078	Sequence 7773, Ap
42	42.4	190	14 US-10-032-585-7773	Sequence 144639,
43	42.4	269	15 US-10-424-599-144639	Sequence 234626,
44	42.4	373	15 US-10-424-599-234626	Sequence 38616, A
45	42.4	373	15 US-10-425-114-38616	Sequence 985, App
46	42.4	456	9 US-09-764-868-985	Sequence 136688,
47	42.4	691	16 US-10-437-963-136688	Sequence 67, Appl
48	42.4	1148	13 US-10-108-605-67	Sequence 173, App
49	42.4	1148	13 US-10-108-605-173	Sequence 275, Appl
50	42.4	1148	13 US-10-108-605-275	Sequence 131308,
51	41.9	1055	16 US-10-437-963-131308	Sequence 223701,
52	41.4	98	17 US-10-425-115-223701	Sequence 265314,
53	41.4	114	17 US-10-425-115-265314	Sequence 261218,
54	41.4	122	15 US-10-424-599-276128	Sequence 266215,
55	41.4	156	17 US-10-425-115-266215	Sequence 93929, Ap
56	41.4	242	17 US-10-739-930-9139	Sequence 232255,
57	41.4	244	15 US-10-424-599-232255	Sequence 169317,
58	41.4	253	16 US-10-437-963-169317	Sequence 40518, A
59	41.4	256	16 US-10-767-701-40518	Sequence 534, App
60	41.4	268	15 US-10-389-566-534	Sequence 195447,
61	41.4	270	15 US-10-424-599-195447	Sequence 313892,
62	41.4	379	17 US-10-425-115-313892	Sequence 47031, A
63	41.4	389	16 US-10-767-701-47031	Sequence 314635,
64	41.4	392	17 US-10-425-115-314635	Sequence 56341, A
65	41.4	395	15 US-10-425-114-56341	Sequence 176, App
66	41.4	487	9 US-09-734-569-176	Sequence 641, App
67	41.4	540	15 US-10-389-566-641	Sequence 620, App
68	41.4	542	15 US-10-389-566-620	Sequence 63666, A
69	41.4	569	15 US-10-282-122A-63666	Sequence 7191, Ap
70	41.4	624	14 US-10-032-585-7191	Sequence 1897, Ap
71	41.4	721	16 US-10-408-765A-1897	Sequence 86, Appl
72	41.4	790	14 US-10-205-823-86	Sequence 17, Appl
73	41.4	790	16 US-10-399-455-17	Sequence 122100,
74	41.4	1090	16 US-10-437-963-122130	Sequence 11, Appl
75	41.4	1464	9 US-09-945-901-11	Sequence 11, Appl
76	41.4	1464	13 US-10-007-747-11	Sequence 11, Appl
77	41.4	1464	13 US-10-038-937-11	Sequence 21, Appl
78	41.4	1464	14 US-10-243-140-2	Sequence 236, Appl
79	41.4	1464	17 US-10-473-127-236	Sequence 237, App
80	41.4	1464	17 US-10-473-127-237	Sequence 238, App
81	41.4	1464	17 US-10-473-127-238	Sequence 240, App
82	41.4	1464	17 US-10-473-127-241	Sequence 241, App
83	41.4	1464	17 US-10-473-127-242	Sequence 242, App
84	41.4	1464	17 US-10-473-127-243	Sequence 243, App
85	41.4	1464	17 US-10-473-127-244	Sequence 244, App

```
87 41 41.4 1464 17 US-10-473-127-245
88 41 41.4 1464 17 US-10-473-127-246
89 40.5 40.9 127 15 US-10-425-114-72863
90 40.5 40.9 278 15 US-10-425-114-65190
91 40.5 40.9 504 15 US-10-425-114-45949
92 40.5 40.9 720 14 US-10-359-493-23409
93 40.5 40.9 798 17 US-10-425-115-20045
94 40 40.4 82 15 US-10-424-599-240616
95 40 40.4 82 15 US-10-424-599-274281
96 40 40.4 117 17 US-10-425-115-220754
97 40 40.4 132 16 US-10-767-701-37853
98 40 40.4 160 15 US-10-424-599-158923
99 40 40.4 216 15 US-10-425-114-41661
100 40 40.4 271 15 US-10-424-599-152589
101 40 40.4 340 15 US-10-282-122A-5310
102 40 40.4 376 15 US-10-424-599-227847
103 40 40.4 379 15 US-10-425-114-36719
104 40 40.4 381 16 US-10-437-963-132654
105 40 40.4 421 9 US-09-420-785A-4
106 40 40.4 421 9 US-09-901-252-15
107 40 40.4 440 15 US-10-425-114-36721
108 40 40.4 445 14 US-10-097-111-289
109 40 40.4 454 14 US-10-029-386-32998
110 40 40.4 463 14 US-10-261-189-10
111 40 40.4 471 17 US-10-425-115-206691
112 40 40.4 477 17 US-10-425-115-207559
113 40 40.4 483 14 US-10-369-493-6691
114 40 40.4 489 15 US-10-389-566-2351
115 40 40.4 489 16 US-10-408-765A-1217
116 40 40.4 490 16 US-10-408-765A-1908
117 40 40.4 498 15 US-10-389-566-2319
118 40 40.4 515 15 US-10-425-114-70123
119 40 40.4 515 15 US-10-425-114-73028
120 40 40.4 528 15 US-10-425-114-51304
121 40 40.4 532 14 US-10-369-493-1926
122 40 40.4 535 15 US-10-606-060A-9
123 40 40.4 548 14 US-10-261-189-11
124 40 40.4 548 15 US-10-389-566-1461
125 40 40.4 658 15 US-10-424-599-167129
126 40 40.4 679 10 US-09-738-269-53
127 40 40.4 679 13 US-10-023-437-53
128 40 40.4 830 14 US-10-211-962-34
129 40 40.4 1310 14 US-10-369-493-3344
130 40 40.4 1536 16 US-10-408-765A-1390
131 40 40.4 1537 17 US-10-643-795A-148
132 40 40.4 1574 17 US-10-128-558-208
133 40 40.4 3365 17 US-10-739-930-6537
134 39.5 39.9 264 15 US-10-425-114-41670
135 39.5 39.9 329 15 US-10-424-599-201446
136 39.5 39.9 732 9 US-09-381-624A-3
137 39.5 39.9 1198 10 US-09-975-719-405
138 39 39.4 43 15 US-10-424-599-147295
139 39 39.4 74 17 US-10-425-115-363170
140 39 39.4 80 17 US-10-425-115-350274
141 39 39.4 91 15 US-10-424-599-216289
142 39 39.4 99 9 US-09-800-729-147
143 39 39.4 99 11 US-09-895-298-72
144 39 39.4 99 11 US-09-833-245-2191
145 39 39.4 99 17 US-10-885-039-72
146 39 39.4 101 17 US-10-425-115-339063
147 39 39.4 108 17 US-10-425-115-204170
148 39 39.4 118 15 US-10-424-599-194265
149 39 39.4 128 15 US-10-424-599-253434
150 39 39.4 227 15 US-10-424-599-162608
```

ALIGNMENTS

```
Sequence 245, App
Sequence 246, App
Sequence 72863, A
Sequence 65190, A
Sequence 45949, A
Sequence 23409, A
Sequence 20045, A
Sequence 240616,
Sequence 274281,
Sequence 220754,
Sequence 37853, A
Sequence 158923,
Sequence 41661, A
Sequence 152589,
Sequence 5310, A
Sequence 227847,
Sequence 36719, A
Sequence 132654,
Sequence 4, Appl1
Sequence 15, Appl1
Sequence 36721, A
Sequence 289, App
Sequence 32998, A
Sequence 10, Appl1
Sequence 206691,
Sequence 207559,
Sequence 6691, Ap
Sequence 2251, Ap
Sequence 1217, Ap
Sequence 1908, Ap
Sequence 2219, Ap
Sequence 70123, A
Sequence 73028, A
Sequence 51304, A
Sequence 1926, Ap
Sequence 9, Appl1
Sequence 11, Appl1
Sequence 1461, Ap
Sequence 167129,
Sequence 53, Appl1
Sequence 53, Appl1
Sequence 34, Appl1
Sequence 3344, Ap
Sequence 1390, Ap
Sequence 148, App
Sequence 208, App
Sequence 6537, Ap
Sequence 41670, A
Sequence 201446,
Sequence 3, Appl1
Sequence 405, App
Sequence 147295,
Sequence 363170,
Sequence 350274,
Sequence 216289,
Sequence 147, App
Sequence 72, Appl1
Sequence 2191, Ap
Sequence 72, Appl1
Sequence 339063,
Sequence 204170,
Sequence 194265,
Sequence 253434,
Sequence 162608,

; GENERAL INFORMATION:
; APPLICANT: Mcintire, Jennifer Jones
; APPLICANT: Umetsu, Dale T.
; APPLICANT: Dekruyff, Rosemarie
; APPLICANT: Kuchroo, Vijay
; APPLICANT: Freeman, Gordon J.
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of
; FILE REFERENCE: STAN-235
; CURRENT APPLICATION NUMBER: US/10/188,012
; PRIOR FILING DATE: 2002-07-01
; SOFTWARE: FastSeq for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 36
; LENGTH: 359
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(360)
; OTHER INFORMATION: TIM-1 allele 1
US-10-188-012-17

Query Match 100.0%; Score 99; DB 14; Length 359;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSDGLMNNNOTQLFLEHS 18
Db 264 SSDGLMNNNOTQLFLEHS 281

RESULT 2
US-10-188-012-19
; Sequence 19, Application US/10188012
; Publication No. US20030124114A1
; GENERAL INFORMATION:
; APPLICANT: Mcintire, Jennifer Jones
; APPLICANT: Umetsu, Dale T.
; APPLICANT: Dekruyff, Rosemarie
; APPLICANT: Kuchroo, Vijay
; APPLICANT: Freeman, Gordon J.
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of
; FILE REFERENCE: STAN-235
; CURRENT APPLICATION NUMBER: US/10/188,012
; PRIOR FILING DATE: 2002-07-01
; SOFTWARE: FastSeq for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 36
; LENGTH: 359
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(359)
; OTHER INFORMATION: TIM-1, allele 2
US-10-188-012-19

Query Match 100.0%; Score 99; DB 14; Length 359;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSDGLMNNNOTQLFLEHS 18
Db 264 SSDGLMNNNOTQLFLEHS 281

RESULT 3
```

```
RESULT 1
US-10-188-012-17
; Sequence 17, Application US/10188012
; Publication No. US20030124114A1
```

US-10-295-027-302
; Sequence 302, Application US/10295027
; Publication No. US2003023250A1
; GENERAL INFORMATION:
; APPLICANT: Afari, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Gineberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezl, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT FILING DATE: US/10/295,027
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 302
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-302

Query Match 100.0%; Score 99; DB 14; Length 359;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSDGLMNNQOTLFLEHS 18
|||
Db 264 SSDGLMNNQOTLFLEHS 281
|||

RESULT 4
US-10-188-832-64
; Sequence 64, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; FILE REFERENCE: 018501-002330US
; CURRENT FILING DATE: US/10/188,832
; PRIOR FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03

; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-832-64

Query Match 100.0%; Score 99; DB 15; Length 359;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSDGLMNNQOTLFLEHS 18
|||
Db 264 SSDGLMNNQOTLFLEHS 281
|||

RESULT 5
US-10-188-012-25
; Sequence 25, Application US/10188012
; Publication No. US2003012411A1
; GENERAL INFORMATION:
; APPLICANT: McIntire, Jennifer Jones
; APPLICANT: Umetsu, Dale T.
; APPLICANT: Dekruyff, Rosemarie
; APPLICANT: Kuchroo, Vijay
; APPLICANT: Freeman, Gordon J.
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of
; FILE REFERENCE: STAN-235
; CURRENT APPLICATION NUMBER: US/10/188,012
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSBQ for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 364
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(364)
; OTHER INFORMATION: T1M-1 allele 5
US-10-188-012-25

Query Match 100.0%; Score 99; DB 14; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.4e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSDGLMNNQOTLFLEHS 18
|||
Db 269 SSDGLMNNQOTLFLEHS 286
|||

RESULT 6
US-10-188-012-27
; Sequence 27, Application US/10188012
; Publication No. US2003012411A1
; GENERAL INFORMATION:
; APPLICANT: McIntire, Jennifer Jones
; APPLICANT: Umetsu, Dale T.
; APPLICANT: Dekruyff, Rosemarie
; APPLICANT: Kuchroo, Vijay
; APPLICANT: Freeman, Gordon J.

```
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of
; FILE REFERENCE: STAN-235
; CURRENT APPLICATION NUMBER: US/10/188,012
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/302,344
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 364
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(364)
; OTHER INFORMATION: TIM-1, allele 6
US-10-188-012-27
```

```
Query Match          100.0%; Score 99; DB 14; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.4e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      1 SSDGLMNNNOTOLFLEHS 18
        |||||:|||||
Db      269 SSDGLMNNNOTOLFLEHS 286
```

```
RESULT 7
US-10-188-012-21
; Sequence 21, Application US/10188012
; Publication No. US20030124114A1
; GENERAL INFORMATION:
; APPLICANT: McIntire, Jennifer Jones
; APPLICANT: Umetsu, Dale T.
; APPLICANT: Dekruyff, Rosemarie
; APPLICANT: Kuchroo, Vijay
; APPLICANT: Freeman, Gordon J.
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of
; FILE REFERENCE: STAN-235
; CURRENT APPLICATION NUMBER: US/10/188,012
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 365
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(365)
; OTHER INFORMATION: TIM-1, allele 3
US-10-188-012-21
```

```
Query Match          100.0%; Score 99; DB 14; Length 365;
Best Local Similarity 100.0%; Pred. No. 2.4e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      1 SSDGLMNNNOTOLFLEHS 18
        |||||:|||||
Db      270 SSDGLMNNNOTOLFLEHS 287
```

```
RESULT 8
US-10-188-012-23
; Sequence 23, Application US/10188012
; Publication No. US20030124114A1
; GENERAL INFORMATION:
; APPLICANT: McIntire, Jennifer Jones
; APPLICANT: Umetsu, Dale T.
```

```
; APPLICANT: Dekruyff, Rosemarie
; APPLICANT: Kuchroo, Vijay
; APPLICANT: Freeman, Gordon J.
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of
; FILE REFERENCE: STAN-235
; CURRENT APPLICATION NUMBER: US/10/188,012
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 359
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(359)
; OTHER INFORMATION: TIM-1, allele 4
US-10-188-012-23
```

```
Query Match          94.9%; Score 94; DB 14; Length 359;
Best Local Similarity 94.4%; Pred. No. 1.5e-06;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      1 SSDGLMNNNOTOLFLEHS 18
        |||||:|||||
Db      264 SSDGLMNNNOTOLFLEHS 281
```

```
RESULT 9
US-09-991-258-13
; Sequence 13, Application US/09991258
; Patent No. US20020141975A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Robert
; APPLICANT: Keith, Paula
; APPLICANT: Dryga, Sergey
; APPLICANT: Caley, Ian
; APPLICANT: Maughan, Maureen
; APPLICANT: Johnston, Robert
; APPLICANT: Davis, Nancy
; APPLICANT: Swansstrom, Ronald
; TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOSES WITH MODIFIED HIV GENES FOR USE I
; FILE REFERENCE: 0113,000103
; CURRENT APPLICATION NUMBER: US/09/991,258
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 09/902,537
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216,995
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 981
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. US20020141975A1 =
US-09-991-258-13
```

```
Query Match          48.5%; Score 48; DB 9; Length 981;
Best Local Similarity 50.0%; Pred. No. 83;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
```

```
Oy      1 SSDGLMNNNOTOLFLE 16
        |||||:|||||
Db      489 SLDHLMNNNOQMFMIQ 504
```

```
RESULT 10
US-10-424-599-175126
; Sequence 175126, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 175126
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_129157C.1.pep
US-10-424-599-175126

Query Match          47.5%; Score 47; DB 15; Length 102;
Best Local Similarity 60.0%; Pred. No. 11;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy      1 SSDGLMNNNOTLFL 15
      ||| : |||
Db      79 STQGLMNNNTHTLL 93

RESULT 11
US-10-369-493-1469
; Sequence 1469, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1469
; LENGTH: 508
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-1469

Query Match          46.5%; Score 46; DB 14; Length 508;
Best Local Similarity 63.6%; Pred. No. 86;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      6 WNNNOTLFL 16
      ||| : |||
Db      160 WNNNASMTLFL 170

RESULT 12
US-10-451-467A-56
; Sequence 56, Application US/10451467A
; Publication No. US20040161840A1
; GENERAL INFORMATION:
; APPLICANT: CONTRERAS, ROLAND HENRI
; APPLICANT: EBERHARDT, INES
; APPLICANT: LUTTEN, WALTER HERMAN MARIA LOUIS
```

```
APPLICANT: REEKMAN, RIEKA JOSEPHINA
; TITLE OF INVENTION: BAX-RESPONSIVE GENES FOR DRUG TARGET IDENTIFICATION IN
; TITLE OF INVENTION: YEAST AND FUNGI
; FILE REFERENCE: USB-1667
; CURRENT APPLICATION NUMBER: US/10/451,467A
; CURRENT FILING DATE: 2003-06-19
; PRIOR APPLICATION NUMBER: EP 00870318.3
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: EP 01870002.1
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: EP 01870003.9
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 732
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56
; LENGTH: 508
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-451-467A-56

Query Match          46.5%; Score 46; DB 16; Length 508;
Best Local Similarity 63.6%; Pred. No. 86;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      6 WNNNOTLFL 16
      ||| : |||
Db      160 WNNNASMTLFL 170

RESULT 13
US-10-389-566-1218
; Sequence 1218, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; APPLICANT: Laurie, Cathy C
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
; FILE REFERENCE: 38-77(52900)D
; CURRENT APPLICATION NUMBER: US/10/389,566
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US 60/365,301
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/391,786
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/392,018
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2459
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1218
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-10-389-566-1218

Query Match          46.5%; Score 46; DB 15; Length 545;
Best Local Similarity 57.1%; Pred. No. 92;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy      3 DGLMNNNOTLFL 16
      ||| : |||
Db      324 DGTWNSXYGPFLE 337

RESULT 14
US-10-437-963-163657
; Sequence 163657, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
```

```

; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 163657
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_62632C.1.pdp
US-10-437-963-163657

Query Match          46.0%; Score 45.5; DB 16; Length 384;
Best Local Similarity 61.1%; Pred. No. 77;
Matches 11; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

Qy      1 SSDG---LMNNNGTQLFL 15
Db      168 SDDGSVKLMNMNQALFL 185

RESULT 15
US-10-437-963-163658
; Sequence 163658, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 163658
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(770)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_62633C.1.pdp
US-10-437-963-163658

Query Match          46.0%; Score 45.5; DB 16; Length 770;
Best Local Similarity 61.1%; Pred. No. 1.6e+02;
Matches 11; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

Qy      1 SSDG---LMNNNGTQLFL 15
Db      554 SDDGSVKLMNMNQALFL 571

RESULT 16
US-10-156-761-9511
; Sequence 9511, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
```

```

; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9511
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9511

Query Match          45.5%; Score 45; DB 14; Length 273;
Best Local Similarity 56.2%; Pred. No. 64;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy      2 SDGLMNNNGTQLFLEH 17
Db      126 SCGEWBERGLFQLFLEH 141

RESULT 17
US-10-282-122A-48522
; Sequence 48522, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zwickind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
```

SEQ ID NO 48522
LENGTH: 687
TYPE: PRT
ORGANISM: Bacteroides fragilis
US-10-282-122A-48522

Query Match 45.5%; Score 45; DB 15; Length 687;
Best Local Similarity 54.5%; Pred. No. 1.7e+02;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

OY 3 SDGLMNNQOTQLFLEH 17
DB 519 EGMMSNDQTEL 529

RESULT 18
US-09-934-070-15
Sequence 15, Application US/09934070
Publication No. US20030092004A1
GENERAL INFORMATION:
APPLICANT: Lipdon, Stuart A.
APPLICANT: Zhang, Dongxian
APPLICANT: Chatterton, Jon E.
APPLICANT: Awobuluyi, Marc
APPLICANT: Sevarino, Kevin A.
TITLE OF INVENTION: EXCITATORY GLYCINE RECEPTORS AND METHODS
FILE REFERENCE: P-LJ 4900
CURRENT APPLICATION NUMBER: US/09/934, 070
CURRENT FILING DATE: 2001-08-20
NUMBER OF SEQ ID NOS: 53
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15
LENGTH: 1464
TYPE: PRT
ORGANISM: Rattus sp.
US-09-934-070-15

Query Match 45.5%; Score 45; DB 10; Length 1464;
Best Local Similarity 47.1%; Pred. No. 3.8e+02;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 1 SDGLMNNQOTQLFLEH 17
DB 1174 NEDGLPNNDOYKLYAKH 1190

RESULT 19
US-09-922-011-1
Sequence 1, Application US/09922011
Publication No. US20030096331A1
GENERAL INFORMATION:
APPLICANT: CIS Biotech, Inc.
APPLICANT: Dambinova, Svetlana
TITLE OF INVENTION: Rapid multiple panel of biomarkers in laboratory blood tests for
FILE REFERENCE: 08805.105001
CURRENT APPLICATION NUMBER: US/09/922, 011
CURRENT FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1464
TYPE: PRT
ORGANISM: homo sapiens
US-09-922-011-1

Query Match 45.5%; Score 45; DB 10; Length 1464;
Best Local Similarity 47.1%; Pred. No. 3.8e+02;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 1 SDGLMNNQOTQLFLEH 17
DB 1174 NEDGLPNNDOYKLYAKH 1190

RESULT 20
US-10-222-772-15
Sequence 15, Application US/10222772
Publication No. US2004003500A1
GENERAL INFORMATION:
APPLICANT: Lipdon, Stuart A.
APPLICANT: Zhang, Dongxian
APPLICANT: Chatterton, Jon E.
APPLICANT: Awobuluyi, Marc
APPLICANT: Sevarino, Kevin A.
TITLE OF INVENTION: EXCITATORY GLYCINE RECEPTORS AND METHODS
FILE REFERENCE: P-LJ 5338
CURRENT APPLICATION NUMBER: US/10/222, 772
CURRENT FILING DATE: 2002-08-16
PRIOR APPLICATION NUMBER: US 09/934, 070
PRIOR FILING DATE: 2001-08-20
NUMBER OF SEQ ID NOS: 62
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15
LENGTH: 1464
TYPE: PRT
ORGANISM: Rattus sp.
US-10-222-772-15

Query Match 45.5%; Score 45; DB 15; Length 1464;
Best Local Similarity 47.1%; Pred. No. 3.8e+02;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 1 SDGLMNNQOTQLFLEH 17
DB 1174 NEDGLPNNDOYKLYAKH 1190

RESULT 21
US-10-739-930-9140
Sequence 9140, Application US/10739930
Publication No. US20040216190A1
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
FILE REFERENCE: 38-21(53377)B
CURRENT APPLICATION NUMBER: US/10/739, 930
CURRENT FILING DATE: 2003-12-18
NUMBER OF SEQ ID NOS: 11088
SEQ ID NO 9140
LENGTH: 243
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: GLYMA-23APR03-C4363_4.P
US-10-739-930-9140

Query Match 44.4%; Score 44; DB 17; Length 243;
Best Local Similarity 52.9%; Pred. No. 82;
Matches 9; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

OY 2 SDGLMNNQOTQLFLEH 18
DB 2 ADGYMNRQA--FLPHS 16

RESULT 22
US-10-424-599-173377
Sequence 173377, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 173377
LENGTH: 268
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_127576C.1.pep
US-10-424-599-173377

Query Match 44.4%; Score 44; DB 15; Length 268;
Best Local Similarity 57.1%; Pred. No. 91;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 LMNNQQLFLEHS 18
Db 155 LMGNFTTPEVNH 168

RESULT 23
US-10-282-122A-57369
Sequence 57369, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zvekind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: EPITRA-034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23/207,727
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22/267,636
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 57369
LENGTH: 432
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-10-282-122A-57369

Query Match 44.4%; Score 44; DB 15; Length 432;
Best Local Similarity 63.6%; Pred. No. 1.5e+02;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 LMNNQQLF 15
Db 72 LMNSKEELFL 82

RESULT 24
US-10-369-493-3911
Sequence 3911, Application US/10369493
Publication No. US2003023675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 3911
LENGTH: 539
TYPE: PRT
ORGANISM: Neurospora crassa
FEATURE:
NAME/KEY: unsure
LOCATION: (1)-(539)
OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-3911

Query Match 44.4%; Score 44; DB 14; Length 539;
Best Local Similarity 43.8%; Pred. No. 1.9e+02;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSDGLMNNQQLFLE 16
Db 202 SNEYAMNNASVIFLD 217

RESULT 25
US-09-826-509-517
Sequence 517, Application US/09826509
Publication No. US20030204073A1
GENERAL INFORMATION:
APPLICANT: Lehmann-Brinsma, Karin
APPLICANT: Liaw, Chen W.
APPLICANT: Lin, I-Lin
TITLE OF INVENTION: No. US20030204073A1-Endogenous, Constitutively Activated Known G
FILE REFERENCE: AREN-207
CURRENT APPLICATION NUMBER: US/09/826,509
CURRENT FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/195,747
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: 09/170,496
PRIOR FILING DATE: 1998-10-13
NUMBER OF SEQ ID NOS: 589
SOFTWARE: PatentIn Version 2.1
SEQ ID NO 517
LENGTH: 590
TYPE: PRT
ORGANISM: Homo sapiens
US-09-826-509-517

Query Match 44.4%; Score 44; DB 10; Length 590;
Best Local Similarity 50.0%; Pred. No. 2.1e+02;

Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
QY 1 SSDGLMNNQTQFLHLS 18
Db 332 SSSDSMNNDAAASLENS 349

RESULT 26
US-10-225-567A-440
Sequence 440, Application US/10225567A
Publication No. US20030113798A1
GENERAL INFORMATION:
APPLICANT: Lifespan Biosciences
APPLICANT: Brown, Joseph P.
APPLICANT: Burner, Gienna C.
APPLICANT: Roush, Christine L.
TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
FILE REFERENCE: 1920-4-4
CURRENT APPLICATION NUMBER: US/10/225,567A
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/257,144
PRIOR FILING DATE: 2000-12-19
NUMBER OF SEQ ID NOS: 2292
SOFTWARE: PatentIn version 3.1
SEQ ID NO 440
LENGTH: 590
TYPE: PRT
ORGANISM: Homo sapiens
US-10-225-567A-440

Query Match 44.4%; Score 44; DB 14; Length 590;
Best Local Similarity 50.0%; Pred. No. 2.1e+02;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
QY 1 SSDGLMNNQTQFLHLS 18
Db 332 SSSDSMNNDAAASLENS 349

RESULT 27
US-10-352-684A-40
Sequence 40, Application US/10352684A
Publication No. US20030215452A1
GENERAL INFORMATION:
APPLICANT: Millenium Pharmaceuticals Inc.
APPLICANT: Carroli, Joseph M.
APPLICANT: Healy, Aileen
APPLICANT: Welch, Nadine S.
APPLICANT: Kelly, Louise M.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
TITLE OF INVENTION: HEMATOLOGICAL DISORDERS USING 131, 148, 199, 12203, 13906,
TITLE OF INVENTION: 15513, 17822, 302, 5677, 194, 14393, 28059, 7366, 12212,
TITLE OF INVENTION: 1981, 261, 12416, 270, 1410, 137, 1871, 13051, 1847, 1849,
FILE REFERENCE: MP102-019PIRKNONIM
CURRENT APPLICATION NUMBER: US/10/352,684A
CURRENT FILING DATE: 2003-01-28
PRIOR APPLICATION NUMBER: US 60/354,333
PRIOR FILING DATE: 2002-02-04
PRIOR APPLICATION NUMBER: US 60/360,258
PRIOR FILING DATE: 2002-02-28
PRIOR APPLICATION NUMBER: US 60/364,476
PRIOR FILING DATE: 2002-03-15
PRIOR APPLICATION NUMBER: US 60/375,626
PRIOR FILING DATE: 2002-04-26
PRIOR APPLICATION NUMBER: US 60/386,494
PRIOR FILING DATE: 2002-06-06
PRIOR APPLICATION NUMBER: US 60/390,965
PRIOR FILING DATE: 2002-06-24
PRIOR APPLICATION NUMBER: US 60/392,480
PRIOR FILING DATE: 2002-06-28
PRIOR APPLICATION NUMBER: US 60/394,128
PRIOR FILING DATE: 2002-07-03

PRIOR APPLICATION NUMBER: US 60/399,783
PRIOR FILING DATE: 2002-07-31
PRIOR APPLICATION NUMBER: US 60/403,221
PRIOR FILING DATE: 2002-08-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 62
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 40
LENGTH: 590
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-352-684A-40

Query Match 44.4%; Score 44; DB 14; Length 590;
Best Local Similarity 50.0%; Pred. No. 2.1e+02;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
QY 1 SSDGLMNNQTQFLHLS 18
Db 332 SSSDSMNNDAAASLENS 349

RESULT 28
US-10-029-009-11
Sequence 11, Application US/10029009
Publication No. US20020164617A1
GENERAL INFORMATION:
APPLICANT: Felsch, Jason S.
APPLICANT: Annis, David Allen
APPLICANT: Kaighatgi, Krishna
APPLICANT: Nash, Huw M.
TITLE OF INVENTION: Affinity Selection-Based Screening of Hydrophobic Proteins
FILE REFERENCE: 11025,173 US2
CURRENT APPLICATION NUMBER: US/10/029,009
CURRENT FILING DATE: 2002-03-28
PRIOR APPLICATION NUMBER: US 60/258,970
PRIOR FILING DATE: 2000-12-29
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 608
TYPE: PRT
ORGANISM: Rat
US-10-029-009-11

Query Match 44.4%; Score 44; DB 13; Length 608;
Best Local Similarity 50.0%; Pred. No. 2.2e+02;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
QY 1 SSDGLMNNQTQFLHLS 18
Db 331 SSSDSMNNDAAASLENS 348

RESULT 29
US-10-029-009-23
Sequence 23, Application US/10029009
Publication No. US20020164617A1
GENERAL INFORMATION:
APPLICANT: Felsch, Jason S.
APPLICANT: Annis, David Allen
APPLICANT: Kaighatgi, Krishna
APPLICANT: Nash, Huw M.
TITLE OF INVENTION: Affinity Selection-Based Screening of Hydrophobic Proteins
FILE REFERENCE: 11025,173 US2
CURRENT APPLICATION NUMBER: US/10/029,009
CURRENT FILING DATE: 2002-03-28
PRIOR APPLICATION NUMBER: US 60/258,970
PRIOR FILING DATE: 2000-12-29
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 23
LENGTH: 626

TYPE: PRT
ORGANISM: Rat
US-10-029-009-23

Query Match 44.4%; Score 44; DB 13; Length 626;
Best Local Similarity 50.0%; Pred. No. 2.2e+02;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 SSDGLMNNQTOLEH 18
DB 347 SSSDSMNNDAAALENS 364

RESULT 30
US-10-369-493-23561
Sequence 23561, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIORITY APPLICATION NUMBER: US 60/360,039
PRIORITY FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 23561
LENGTH: 1042
TYPE: PRT
ORGANISM: Escherichia coli
US-10-369-493-23561

Query Match 44.4%; Score 44; DB 14; Length 1042;
Best Local Similarity 46.7%; Pred. No. 3.8e+02;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 DGLMNNQTOLEH 17
DB 788 EGLMQPNHLQIMQH 802

RESULT 31
US-10-424-599-282496
Sequence 282496, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 282496
LENGTH: 48
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_97115C.1 pep
US-10-424-599-282496

Query Match 43.4%; Score 43; DB 15; Length 48;
Best Local Similarity 58.3%; Pred. No. 21;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 LMNNQTOLEH 16
DB 7 MMRKQTKSFLE 18

RESULT 32
US-10-424-599-174021
Sequence 174021, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 174021
LENGTH: 100
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(100)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_128159C.1 pep
US-10-424-599-174021

Query Match 43.4%; Score 43; DB 15; Length 100;
Best Local Similarity 50.0%; Pred. No. 46;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 5 LMNNQTOLEH 16
DB 84 LMNNHQIEFFVD 95

RESULT 33
US-10-425-115-312882
Sequence 312882, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 312882
LENGTH: 116
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_48408C.1 pep
US-10-425-115-312882

Query Match 43.4%; Score 43; DB 17; Length 116;
Best Local Similarity 42.9%; Pred. No. 54;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 5 LMNNQTOLEH 18
DB 22 LMNSNHTGILIKKN 35

```
RESULT 34
US-10-221-625-65
; Sequence 65, Application US/10221625
; Publication No. US20040033942A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: YUE, Henry
; APPLICANT: LAL, Preeti
; APPLICANT: LU, Dyrng Alina M.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BANDMAN, Olga
; APPLICANT: TANG, Y. Tom
; APPLICANT: MATHUR, Preetee
; APPLICANT: SHAH, Purvi
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: REDDY, Roopa
; TITLE OF INVENTION: TRANSCRIPTION FACTORS
; FILE REFERENCE: PP-0761 PCT
; CURRENT APPLICATION NUMBER: US/10/221,625
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PERL Program
; SEQ ID NO 65
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040033942A1 2646274CD1
US-10-221-625-65

Query Match          43.4%; Score 43; DB 15; Length 233;
Best Local Similarity 47.1%; Pred. No. 1.1e+02;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Oy 1 SSDGLMNNOTQPLEH 17
Db 75 STEGLFNNGFAEAFLEH 91

RESULT 35
US-10-437-963-176868
; Sequence 176868, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 176868
; LENGTH: 554
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_74577C.1.pep
US-10-437-963-176868

Query Match          43.4%; Score 43; DB 16; Length 554;
Best Local Similarity 50.0%; Pred. No. 2.8e+02;
Matches 10; Conservative 2; Mismatches 4; Indels 4; Gaps 1;
```

```
Oy 1 SSDGLMNNOTQPLEH 16
Db 496 ASDGLMDVTSANRAAQVLE 515

RESULT 36
US-10-094-466-8
; Sequence 8, Application US/10094466
; Publication No. US20030203363A1
; GENERAL INFORMATION:
; APPLICANT: Spylek et al.
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM
; TITLE OF INVENTION: AND METHODS OF USING
; FILE REFERENCE: 21402-290D
; CURRENT APPLICATION NUMBER: US/10/094,466
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/288,148
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/275,235
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/338,375
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/275,579
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/335,302
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/275,601
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/276,000
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/277,338
; PRIOR FILING DATE: 2001-03-20
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patin 2.1
; SEQ ID NO 8
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-466-8

Query Match          43.4%; Score 43; DB 14; Length 561;
Best Local Similarity 47.1%; Pred. No. 2.8e+02;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Oy 1 SSDGLMNNOTQPLEH 17
Db 279 STEGLFNNGFAEAFLEH 295

RESULT 37
US-10-437-963-174682
; Sequence 174682, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
```

```
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 174682
; LENGTH: 684
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_725C.1.pep
US-10-437-963-174682
```

```
Query Match          43.4%; Score 43; DB 16; Length 684;
Best Local Similarity 63.6%; Pred. No. 3.5e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      5 LMNNNQTQLFLL 15
         ||:||:|||
Db       431 LMSNNKELFL 441
```

```
RESULT 38
US-10-156-761-9394
; Sequence 9394, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9394
; LENGTH: 973
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9394
```

```
Query Match          43.4%; Score 43; DB 14; Length 973;
Best Local Similarity 63.6%; Pred. No. 5.1e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      3 DGLMNNNQTQL 13
         ||:||:|||
Db       872 DGMNNNYQDSL 882
```

```
RESULT 39
US-10-424-599-265756
; Sequence 265756, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 265756
; LENGTH: 42
; TYPE: PRT
```

```
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_81999C.1.pep
US-10-424-599-265756
```

```
Query Match          42.4%; Score 42; DB 15; Length 42;
Best Local Similarity 47.1%; Pred. No. 27;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
```

```
QY      2 SDGLMNNNQTQLFLEHS 18
         ||:||:|||
Db       17 STLMNNWVDFLEHN 33
```

```
RESULT 40
US-10-425-115-367358
; Sequence 367358, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 367358
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_98200C.1.pep
US-10-425-115-367358
```

```
Query Match          42.4%; Score 42; DB 17; Length 50;
Best Local Similarity 42.9%; Pred. No. 32;
Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      5 LMNNNQTQLFLEHS 18
         ||:||:|||
Db       9 VMSNHRVGRFLQHS 22
```

```
Search completed: December 30, 2004, 16:25:32
Job time : 147 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 16:14:08 ; Search time 152 Seconds

(without alignments)
42.481 Million cell updates/sec

Title: US-10-718-321-1

Sequence: 1 SSDGLMNNQTOLELHHS 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729239 residues

Total number of hits satisfying chosen parameters: 686242

Minimum DB seq length: 0

Maximum DB seq length: 20

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : A_Geneseq_23Sep04:*
1: geneseq1980s:*
2: geneseq1990s:*
3: geneseq2000s:*
4: geneseq2001s:*
5: geneseq2002s:*
6: geneseq2003as:*
7: geneseq2003bs:*
8: geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	99	100.0	18 6	AAO26687 Human KIM
2	99	100.0	18 6	AAO26673 Monoclonal
3	53	53.5	18 6	AAO26686 Human KIM
4	46.5	46.5	18 6	AAO26688 Human KIM
5	35.4	18 8	ADH35129	Adh35129 Complement
6	35.4	19 8	ADH35000	Adh35000 N-linked
7	34.3	12 7	ADC36109	Adc36109 Chemokine
8	34.3	19 6	ABU13471	Abu13471 Zona pell
9	34.3	20 6	ABP83017	Abp83017 G protein
10	33.3	18 8	ABE11136	ABE11136 Antigenic
11	33.3	20 5	ABG80887	Abg80887 Human ebl
12	33.3	20 6	ABO14120	Abol14120 Novel hum
13	33.3	20 8	ADG78515	Adg78515 Human sec
14	33.3	20 8	ADN60818	Adn60818 Human sec
15	32.3	12 2	AKR91293	AKr91293 Anti-Idio
16	32.3	18 2	AAW09338	AAw09338 FIV princ
17	32.3	19 2	AAW07984	AAw07984 gp41 pepc
18	31.3	15 4	AAAG64600	AAag64600 HSV-4 BFL
19	31.3	15 8	ADJ38603	Adj38603 Hsv-4 BFL
20	31.3	15 8	ADL25941	Adl25941 Synthetic
21	31.3	15 8	ADL25831	Adl25831 Synthetic
22	31.3	15 8	ADL25933	Adl25933 Synthetic
23	31.3	15 8	ADL25932	Adl25932 Synthetic
24	31.3	15 8	ADL25940	Adl25940 Synthetic
25	31.3	17 5	AAW50838	AAw50838 HCV E2 re

26	31	31.3	19 7	ABW01700 Toxoplas
27	31	31.3	19 7	ADOS9249 Toxoplas
28	31	31.3	20 6	ADA26471 Isomaltas
29	31	31.3	20 7	ADH80940 HIV hyper
30	31	31.3	20 8	ADM29206 Hepatitis
31	31	31.3	20 8	ABO57864 Human gen
32	30	30.3	10 7	ADE78906 Interleuk
33	30	30.3	12 4	AAE10548 llama spe
34	30	30.3	12 5	ABH80837 Hepatin b
35	30	30.3	14 4	ABW73315 Human PC-
36	30	30.3	14 5	ABU00795 B lymphoc
37	30	30.3	14 5	ABG33656 B lymphoc
38	30	30.3	15 2	AAH85392 HTLV-IIIB
39	30	30.3	15 2	AAW19031 SP-15 reg
40	30	30.3	15 8	ADL25934 Synthetic
41	30	30.3	17 2	AAH60682 Synthetic
42	30	30.3	17 2	AAW73262 SP-10 11k
43	30	30.3	17 5	ABG30549 Alpha-180
44	30	30.3	18 3	AAV51197 Human IL-
45	30	30.3	20 2	AAH84434 Hepatitis
46	30	30.3	20 2	AAH84511 Hepatitis
47	30	30.3	20 2	AAH84433 Hepatitis
48	30	30.3	20 2	AAW08062 HIV pepti
49	30	30.3	20 2	AAH91005 HCV E2 pe
50	30	30.3	20 2	AAH90996 HCV E2 pe
51	30	30.3	20 5	AAO18702 Hepatitis
52	30	30.3	20 5	AAO18711 Hepatitis
53	30	30.3	20 5	AAE17442 Dengue vi
54	30	30.3	20 5	ABG80897 Human pel
55	30	30.3	20 5	ABG80898 Human pel
56	30	30.3	20 5	ABG80893 Human pel
57	30	30.3	20 5	ABG80895 Human pel
58	30	30.3	20 5	ABG80896 Human pel
59	30	30.3	20 5	ABG80894 Human pel
60	30	30.3	20 7	ADU55590 Hepatitis
61	30	30.3	20 7	ADU55581 Hepatitis
62	30	30.3	20 8	ADN36390 HIV gag e
63	30	30.3	20 8	ADP71163 HCV E2 pr
64	30	30.3	20 8	ADP71172 HCV E2 pr
65	29	29.3	7 2	AAH72562 Rat manno
66	29	29.3	7 2	AAV41833 Rat manno
67	29	29.3	7 2	AAW95243 Rat manno
68	29	29.3	7 3	AAH68382 Mannose b
69	29	29.3	7 4	AAH66256 MBP C-ter
70	29	29.3	7 5	AAU70373 Mouse hea
71	29	29.3	9 7	ADM18272 C trachom
72	29	29.3	10 6	ABR26125 Human can
73	29	29.3	10 6	ABR26125 Human can
74	29	29.3	10 6	ABR26705 Human can
75	29	29.3	10 6	ABR26946 Human can
76	29	29.3	13 3	AAV51868 Pancreat
77	29	29.3	14 8	ADOL10179 Human int
78	29	29.3	15 4	ABP24509 HIV DR su
79	29	29.3	15 5	ABG62729 Eubacteri
80	29	29.3	15 5	ABG62736 Eubacteri
81	29	29.3	15 5	ABG62737 Eubacteri
82	29	29.3	15 5	ABG62735 Eubacteri
83	29	29.3	15 5	ABG80878 Rat MBP m
84	29	29.3	15 5	ABG80878 Rat MBP m
85	29	29.3	15 5	ABG80874 Rat MBP m
86	29	29.3	15 5	ABG80873 Rat MBP m
87	29	29.3	15 5	ABG80875 Rat MBP m
88	29	29.3	15 5	ABR38103 Human can
89	29	29.3	15 6	ABR38158 Human can
90	29	29.3	15 6	ABR38239 Human can
91	29	29.3	15 6	ABR38212 Human can
92	29	29.3	15 8	ADH10427 S. pneumo
93	29	29.3	15 8	ADJ38601 Hsv-4 BFL
94	29	29.3	15 8	ADL25935 Synthetic
95	29	29.3	15 8	ADL26374 Synthetic
96	29	29.3	16 7	ADD94226 Mouse HUI
97	29	29.3	16 8	ADM18969 HLA-DR bo
98	29	29.3	17 2	AAW48681 Amino aci

99	29	29.3	17	4	AAU72451	Aau72451	MAGE-1/3-
100	29	29.3	19	2	AAU24864	Aa224864	Sequence
101	29	29.3	20	6	ABR52459	Ab252459	Group A s
102	29	29.3	20	8	ADH37233	Adh37233	Epretein B
103	29	29.3	20	8	ADH37234	Adh37234	Epretein B
104	29	29.3	20	8	ADH29217	Adm29217	Hepatitis
105	28	28.8	15	8	ADP26565	Adp26565	Plasmodiu
106	28	28.3	8	2	AAW15980	Aaw15980	Interleuk
107	28	28.3	8	2	AAW68619	Aaw68619	Random pe
108	28	28.3	8	2	AAW58094	Aaw58094	Peptide S
109	28	28.3	8	3	AAV09742	Aav09742	Interleuk
110	28	28.3	8	3	ABH17545	Abh17545	VIP-mimic
111	28	28.3	8	5	ABB73147	Abb73147	VIP mimet
112	28	28.3	8	7	ADJ73301	Adj73301	Vasocactiv
113	28	28.3	8	8	ADJ52935	Adj52935	CH1 delet
114	28	28.3	8	8	ADJ51896	Adj51896	CH1 delet
115	28	28.3	9	2	AAW39838	Aaw39838	Light cha
116	28	28.3	9	2	ABR11854	Ab11854	Human can
117	28	28.3	9	6	ABR12626	Ab112626	Human can
118	28	28.3	9	6	ABR13035	Ab113035	Human can
119	28	28.3	9	6	ABR12250	Ab112250	Human can
120	28	28.3	10	6	ABR12361	Ab112361	Human can
121	28	28.3	10	6	ABR11883	Ab111883	Human can
122	28	28.3	10	6	ABR12318	Ab112318	Human can
123	28	28.3	10	6	ABR12482	Ab112482	Human can
124	28	28.3	10	6	ABR12695	Ab112695	Human can
125	28	28.3	11	2	ABR11909	Ab111909	Human can
126	28	28.3	11	2	AAW54750	Aaw54750	Ad5 fibre
127	28	28.3	12	2	AAW07638	Aaw07638	P. suis 1
128	28	28.3	12	2	AAW68895	Aaw68895	Peptide o
129	28	28.3	12	2	AAW60549	Aaw60549	Oligopept
130	28	28.3	12	2	AAW58091	Aaw58091	Peptide S
131	28	28.3	12	4	AAW58090	Aaw58090	Peptide S
132	28	28.3	12	4	AAW50187	Aaw50187	Mycobacte
133	28	28.3	13	3	AAV94566	Aav94566	Vasocactiv
134	28	28.3	13	5	AAE27227	Aae27227	Human int
135	28	28.3	13	5	ADG66480	Adg66480	B. amylo1
136	28	28.3	13	5	ADG65580	Adg65580	Human IL-
137	28	28.3	13	6	AAE33009	Aae33009	Human int
138	28	28.3	13	7	ADJ89122	Adj89122	Human her
139	28	28.3	14	2	AAE69361	Aae69361	Stearyl-
140	28	28.3	14	2	AAV94563	Aav94563	Vasocactiv
141	28	28.3	14	4	AAJ98115	Aaj98115	Human SNP
142	28	28.3	14	5	ABG34903	Abg34903	Human G-P
143	28	28.3	14	5	ABJ00739	Abj00739	B lymphoc
144	28	28.3	14	5	ABJ00775	Abj00775	B lymphoc
145	28	28.3	14	5	ABJ00765	Abj00765	B lymphoc
146	28	28.3	14	5	ABG33636	Abg33636	B lymphoc
147	28	28.3	14	5	ABG33626	Abg33626	B lymphoc
148	28	28.3	14	5	ABG33600	Abg33600	B lymphoc
149	28	28.3	14	6	ADA03332	Ada03332	Angiotens
150	28	28.3	14	6	ABR56702	Ab256702	Angiotens

ALIGNMENTS

RESULT 1
AAO26687
ID AAO26687 standard; peptide; 18 AA.

XX AAO26687;

XX 20-MAR-2003 (first entry)

XX Human KIM-1 mucin domain related peptide, SEQ ID No 15.

XX Cytostatic; gene therapy; antibody; antigen; antigen-binding;

XX proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;

XX renal disease; injury; renal cancer.

XX Synthetic.

PN	WO200298920-A1.
XX	12-DEC-2002.
XX	31-MAY-2002; 2002MO-US017402.
PF	01-JUN-2001; 2001US-0295449P.
XX	04-JUN-2001; 2001US-0295907P.
PR	(BIOJ) BIOGEN INC.
XX	(GENO) GEN HOSPITAL CORP.
PA	Bailly V, Bonventre J;
XX	WPI; 2003-156845/15.
DR	
XX	New antibody, antibody derivative or antigen-binding polypeptide that
PT	inhibits proteolytic release of a soluble kidney injury molecule-1
PT	polypeptide, useful for treating or preventing renal disease or injury,
PT	e.g. renal cancer.
XX	
PS	Disclosure; Fig 1A; 42pp; English.
XX	
CC	The invention relates to a novel antibody, antibody derivative or antigen
CC	-binding polypeptide that inhibits proteolytic release of a soluble
CC	kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
CC	The antibody, antibody derivative or antigen-binding polypeptide is
CC	useful for treating or preventing renal disease or injury, e.g. renal
CC	cancer. The antibody is also useful for inhibiting shedding of the KIM-1
CC	polypeptide. This sequence represents a human KIM-1 mucin domain related
CC	peptide of the invention
XX	
SQ	Sequence 18 AA;

Query Match 100.0%; Score 99; DB 6; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.7e-08;
Matches 18; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSDGLMNNNOTFLFHS 18
Db 1 SSDGLMNNNOTFLFHS 18

RESULT 2
AAO26673
ID AAO26673 standard; peptide; 18 AA.

XX AAO26673;

XX 20-MAR-2003 (first entry)

XX Monoclonal antibody binding epitope, SEQ ID No 1.

XX Cytostatic; gene therapy; antibody; antigen; antigen-binding;

XX proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;

XX renal disease; injury; renal cancer; binding epitope.

XX Homo sapiens.

XX WO200298920-A1.

XX 12-DEC-2002.

XX 31-MAY-2002; 2002MO-US017402.

XX 01-JUN-2001; 2001US-0295449P.

XX 04-JUN-2001; 2001US-0295907P.

XX (BIOJ) BIOGEN INC.

XX (GENO) GEN HOSPITAL CORP.

XX Bailly V, Bonventre J;

DR WPI; 2003-156845/15.
 XX New antibody, antibody derivative or antigen-binding polypeptide that
 PT inhibits proteolytic release of a soluble kidney injury molecule-1
 PT polypeptide, useful for treating or preventing renal disease or injury,
 PT e.g. renal cancer.
 XX
 PS Claim 4; Page 25; 42pp; English.
 XX
 CC The invention relates to a novel antibody, antibody derivative or antigen
 CC -binding polypeptide that inhibits proteolytic release of a soluble
 CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
 CC The antibody, antibody derivative or antigen-binding polypeptide is
 CC useful for treating or preventing renal disease or injury, e.g. renal
 CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
 CC polypeptide. This sequence represents a binding epitope of the invention
 XX
 SQ Sequence 18 AA;
 Query Match 100.0%; Score 99; DB 6; Length 18;
 Best Local Similarity 100.0%; Pred. No. 4.7e-06;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 SSDGLMNNNOTFLFLEHS 18
 Db 1 SSDGLMNNNOTFLFLEHS 18
 RESULT 3
 AAO26686
 ID AAO26686 standard; peptide; 18 AA.
 AC AAO26686;
 XX
 DT 20-MAR-2003 (first entry)
 XX
 DE Human KIM-1 mucin domain related peptide, SEQ ID No 14.
 XX
 KM Cytostatic; gene therapy; antibody; antigen; antigen-binding;
 KM proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
 KM renal disease; injury; renal cancer.
 XX
 OS Synthetic.
 XX
 PI Baillly V, Bonventure J;
 XX
 DR WPI; 2003-156845/15.
 XX
 PT New antibody, antibody derivative or antigen-binding polypeptide that
 PT inhibits proteolytic release of a soluble kidney injury molecule-1
 PT polypeptide, useful for treating or preventing renal disease or injury,
 PT e.g. renal cancer.
 XX
 PS Disclosure; Fig 1A; 42pp; English.
 XX
 CC The invention relates to a novel antibody, antibody derivative or antigen
 CC -binding polypeptide that inhibits proteolytic release of a soluble
 CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
 CC The antibody, antibody derivative or antigen-binding polypeptide is
 CC useful for treating or preventing renal disease or injury, e.g. renal
 CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1

CC polypeptide. This sequence represents a human KIM-1 mucin domain related
 CC peptide of the invention
 XX
 SQ Sequence 18 AA;
 Query Match 53.5%; Score 53; DB 6; Length 18;
 Best Local Similarity 100.0%; Pred. No. 0.45;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 SSDGLMNNN 9
 Db 10 SSDGLMNNN 18
 RESULT 4
 AAO26688
 ID AAO26688 standard; peptide; 18 AA.
 AC AAO26688;
 XX
 DT 20-MAR-2003 (first entry)
 XX
 DE Human KIM-1 mucin domain related peptide, SEQ ID No 16.
 XX
 KM Cytostatic; gene therapy; antibody; antigen; antigen-binding;
 KM proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
 KM renal disease; injury; renal cancer.
 XX
 OS Synthetic.
 XX
 PI Baillly V, Bonventure J;
 XX
 DR WPI; 2003-156845/15.
 XX
 PT New antibody, antibody derivative or antigen-binding polypeptide that
 PT inhibits proteolytic release of a soluble kidney injury molecule-1
 PT polypeptide, useful for treating or preventing renal disease or injury,
 PT e.g. renal cancer.
 XX
 PS Disclosure; Fig 1A; 42pp; English.
 XX
 CC The invention relates to a novel antibody, antibody derivative or antigen
 CC -binding polypeptide that inhibits proteolytic release of a soluble
 CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
 CC The antibody, antibody derivative or antigen-binding polypeptide is
 CC useful for treating or preventing renal disease or injury, e.g. renal
 CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
 CC polypeptide. This sequence represents a human KIM-1 mucin domain related
 CC peptide of the invention
 XX
 SQ Sequence 18 AA;
 Query Match 46.5%; Score 46; DB 6; Length 18;
 Best Local Similarity 100.0%; Pred. No. 5.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 10 QTOFLFLEHS 18
 Db 1 QTOFLFLEHS 9

RESULT 5
ADH35129
ID ADH35129 standard; peptide; 18 AA.
XX
AC ADH35129;
XX
DT 11-MAR-2004 (first entry)
XX
DE Complement c4 glycopeptide SEQ ID NO:188.
XX
KW identification; quantification; glycopolypeptide; glycopeptide;
KW diagnostic marker; solid support; glycosylated polypeptide;
KW differential glycosylation; cancer; glycoprotein; characterisation;
KW diagnostic biomarker; immunotherapy; drug development; toxicology;
KW drug targeting.
XX
OS Synthetic.
OS Homo sapiens.
PN W02003102018-A2.
PD 11-DEC-2003.
XX
DE 03-JUN-2003; 2003WO-US017635.
XX
PR 03-JUN-2002; 2002US-0385707P.
PR 09-MAY-2003; 2003US-0469361P.
XX
PA (SYST-) INST SYSTEMS BIOLOGY.
XX
PI Aebersold RH, Zhang H;
XX
DR WPI; 2004-081931/08.
XX
PT Identifying and quantifying glycopolypeptides in sample by derivatizing
PT glycopolypeptides, immobilizing and cleaving derivatized
PT glycopolypeptides, identifying and quantifying released glycopeptide
PT fragments.
XX
XX Example 6; SEQ ID NO 188; 162pp; English.
XX
CC The present invention describes a method (M1) for identifying and
CC quantifying glycopolypeptides (I) in a sample. M1 comprises derivatising
CC (I), immobilising the derivatised (I), cleaving the immobilised (I),
CC labelling the immobilised glycopeptide fragments (II) with an isotope
CC tag, releasing (II) from the solid support, analysing the released (II)
CC using mass spectrometry, identifying a released (II) and quantifying the
CC amount of (II). The method alternatively comprises: (a) immobilising (I)
CC to a solid support; (b) cleaving the immobilised (I); (c) labelling the
CC immobilised glycopeptides (III) with an isotope tag; and/or (c) releasing
CC (III) from the solid support; and (d) analysing the released (III). Also
CC described: (i) identifying (M2) diagnostic marker for a disease.
CC comprising: (a) immobilising (I) from a test sample or control sample to
CC a first solid support or second solid support respectively; (b) cleaving
CC the immobilised (I); (c) labelling the immobilised (III) on the first and
CC second supports with differential isotope tags on the respective supports
CC (d) releasing (III) from the solid support; (e) analysing (III); and
CC (f) identifying one or more glycosylated polypeptides having differential
CC glycosylation between the test sample and the control sample; and (2) a
CC kit comprising a hydrazone resin, periodate and a pair of differentially
CC labelled isotope tags. (M1) is useful for identifying and quantifying (I)
CC in a sample and for identifying diagnostic marker for a disease such as
CC cancer. (M1) is useful for determining the sites of glycosylation,
CC determining changes in the abundance of glycoproteins and changes in the
CC state of glycosylation at individual glycosylation sites on those
CC glycoproteins that occur in response to perturbations of biological
CC systems and organisms in health and disease. (M1) is also useful for
CC purifying glycosylated proteins or peptides and for isolating and
CC identifying glycoproteins from the cell membrane or body fluids to
CC determine specific glycoprotein changes related to certain disease states
CC or cancer. (M1) is also useful for identifying, selecting and
CC characterising subgroups of carbohydrates and characterisation of
CC diagnostic biomarkers, immunotherapy and to evaluate the effectiveness of

CC drugs during development, optimal dosing, toxicology, drug targeting and
CC related therapeutic applications. The present sequence represents a
CC glycopeptide which is used in an example from the present invention.
XX
SQ Sequence 18 AA;
XX
Query Match 35.4%; Score 35; DB 8; Length 18;
Best Local Similarity 63.6%; Pred. No. 2,4e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Db 3 SDGSMNSSTQ 13
XX
RESULT 6
ADH35000
ID ADH35000 standard; peptide; 19 AA.
XX
AC ADH35000;
XX
DT 11-MAR-2004 (first entry)
XX
DE N-linked glycopeptide #19.
XX
KW identification; quantification; glycopolypeptide; glycopeptide;
KW diagnostic marker; solid support; glycosylated polypeptide;
KW differential glycosylation; cancer; glycoprotein; characterisation;
KW diagnostic biomarker; immunotherapy; drug development; toxicology;
KW drug targeting.
XX
OS Synthetic.
OS Homo sapiens.
PN W02003102018-A2.
PD 11-DEC-2003.
XX
DE 03-JUN-2003; 2003WO-US017635.
XX
PR 03-JUN-2002; 2002US-0385707P.
PR 09-MAY-2003; 2003US-0469361P.
XX
PA (SYST-) INST SYSTEMS BIOLOGY.
XX
PI Aebersold RH, Zhang H;
XX
DR WPI; 2004-081931/08.
XX
PT Identifying and quantifying glycopolypeptides in sample by derivatizing
PT glycopolypeptides, immobilizing and cleaving derivatized
PT glycopolypeptides, identifying and quantifying released glycopeptide
PT fragments.
XX
XX Example 15; Fig 28; 162pp; English.
XX
CC The present invention describes a method (M1) for identifying and
CC quantifying glycopolypeptides (I) in a sample. M1 comprises derivatising
CC (I), immobilising the derivatised (I), cleaving the immobilised (I),
CC labelling the immobilised glycopeptide fragments (II) with an isotope
CC tag, releasing (II) from the solid support, analysing the released (II)
CC using mass spectrometry, identifying a released (II) and quantifying the
CC amount of (II). The method alternatively comprises: (a) immobilising (I)
CC to a solid support; (b) cleaving the immobilised (I); (c) labelling the
CC immobilised glycopeptides (III) with an isotope tag; and/or (c) releasing
CC (III) from the solid support; and (d) analysing the released (III). Also
CC described: (i) identifying (M2) diagnostic marker for a disease,
CC comprising: (a) immobilising (I) from a test sample or control sample to
CC a first solid support or second solid support respectively; (b) cleaving
CC the immobilised (I); (c) labelling the immobilised (III) on the first and
CC second supports with differential isotope tags on the respective supports
CC (d) releasing (III) from the solid support; (e) analysing (III); and
CC (f) identifying one or more glycosylated polypeptides having differential
CC glycosylation between the test sample and the control sample; and (2) a

CC kit comprising a hydrazone resin, periodate and a pair of differentially
 CC labelled isotope tags. (M1) is useful for identifying and quantifying (I)
 CC in a sample and for identifying diagnostic marker for a disease such as
 CC cancer. (M1) is useful for determining the sites of glycosylation,
 CC determining changes in the abundance of glycoproteins and changes in the
 CC state of glycosylation at individual glycosylation sites on those
 CC glycoproteins that occur in response to perturbations of biological
 CC systems and organisms in health and disease. (M1) is also useful for
 CC purifying glycosylated proteins or peptides and for isolating and
 CC identifying glycoproteins from the cell membrane or body fluids to
 CC determine specific glycoprotein changes related to certain disease states
 CC or cancer. (M1) is also useful for identifying, selecting and
 CC characterising subgroups of carbohydrates and characterisation of
 CC diagnostic biomarkers, immunotherapy and to evaluate the effectiveness of
 CC drugs during development, optimal dosing, toxicology, drug targeting and
 CC related therapeutic applications. The present sequence represents an N-
 CC linked glycopeptide which is used in an example from the present
 CC invention.

CC XX Sequence 19 AA;

SO Query Match 35.4%; Score 35; DB 8; Length 19;
 Best Local Similarity 63.6%; Pred. No. 2.6e+02;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 SDGLMNNQTO 12
 ||| : |||
 Db 3 SDGLSNSTO 13

RESULT 7

ID ADC36109 standard; peptide; 12 AA.

XX AC ADC36109;

DT 18-DEC-2003 (first entry)

DE Chemokine binding peptide BKT-P143.

XX peptide chemokine modulator; antiinflammatory; antiallergic;
 KW immunosuppressive; antidiabetic; antirheumatic; dermatological;
 KW antiarthritic; antibacterial; antiproliferative; antineoplastic;
 KW antiarteriosclerotic; hypotensive; neuroprotective; virucide; vasotropic;
 KW cytostatic; inflammation; allergy; immune response; autoimmune reaction;
 KW rheumatoid arthritis; systemic lupus erythematosus; multiple sclerosis;
 KW allograft rejection; diabetes; sepsis; cancer; malignant cell growth;
 KW acne; infection; arthritis; colitis; psoriasis; atherosclerosis;
 KW hypertension; reperfusion ischaemia.

XX OS Synthetic.

XX PN WO200307259-A2.

XX PD 04-SEP-2003.

XX PF 27-FEB-2003; 2003WO-IL000155.

XX PR 28-FEB-2002; 2002US-0359995P.

XX PA (BIOK-) BIOKINE THERAPEUTICS LTD.

XX PI Peled A, Eizenberg O, Valzel-Ohayon D;

XX DR WPI; 2003-671869/63.

XX New peptidic chemokine modulator, useful for preparing a composition for
 PT treating a disease modulated through and/or caused by binding of a
 PT chemokine to a chemokine receptor, e.g., inflammation, diabetes, sepsis
 PT or cancer.

XX Example 1; Page 29; 43pp; English.

XX

CC The present invention describes a peptidic chemokine modulator (I) for
 CC modulating a biological effect of a chemokine. (I) comprises a molecule
 CC consisting of: (a) the amino acids His, Ser, Ala, Leu, Ile, Lys, Arg, Thr
 CC and Pro, and features at least 2 histidines spread along the molecule,
 CC where the molecule features an overall positive charge (family 1); or (b)
 CC the amino acids His, Pro, Thr, Leu, Arg and Tyr and features at least two
 CC neighbouring histidines, where the molecule features an overall positive
 CC charge (family 2). Also described: (1) a composition for treating a
 CC condition involving abnormal cell migration in a subject; (2) a method
 CC for treating a disease modulated through and/or caused by binding of a
 CC chemokine to a chemokine receptor in a subject; (3) an antibody for
 CC binding to a chemokine-binding receptor that recognises at least a
 CC portion of a chemokine-binding receptor or the peptide; (4) a vaccine
 CC formed with the antibody; and (5) a method for producing an antibody. (I)
 CC has antiinflammatory, antiallergic, immunosuppressive, antidiabetic,
 CC antineumatic, dermatological, antiarthritic, antibacterial,
 CC antiproliferative, antineoplastic, antirheumatic, antidiabetic, hypotensive,
 CC neuroprotective, virucide, vasotropic and cytostatic activities. The
 CC peptidic chemokine modulator is useful for preparing a composition for
 CC treating a disease modulated through and/or caused by binding of a
 CC chemokine to a chemokine receptor, comprising inflammation (primary or
 CC secondary), allergy, a non-optimal immune response, an autoimmune
 CC reaction (including rheumatoid arthritis, systemic lupus erythematosus,
 CC multiple sclerosis and others), allograft rejection, diabetes, sepsis,
 CC cancer and any type of malignant cell growth, acne and chronic bacterial
 CC and viral infections, arthritis, colitis, psoriasis, atherosclerosis,
 CC hypertension or reperfusion ischaemia. The present sequence represents a
 CC chemokine binding peptide, which is used in an example from the present
 CC invention.

SO Sequence 12 AA;

Query Match 34.3%; Score 34; DB 7; Length 12;
 Best Local Similarity 71.4%; Pred. No. 2.2e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 6 WNNNQTO 12
 ||| : |||
 Db 2 WNSNSTO 8

RESULT 8

ID ABU13471 standard; peptide; 19 AA.

XX AC ABU13471;

DT 21-FEB-2003 (first entry)

DE Zona pellucida protein (ZP1) peptide #66.

XX Human; zona pellucida; immunogen; ZP1; sperm; contraceptive; vaccine.

XX OS Homo sapiens.

XX PN US6455041-B1.

XX PD 24-SEP-2002.

XX PF 17-NOV-1999; 99US-00441502.

XX PR 17-NOV-1998; 98US-0108822P.

XX PA (DUNB/) DUNBAR B S.

XX PI Dunbar BS;

XX DR WPI; 2003-089265/08.

XX Inducing anti-zona pellucida protein antibodies in a mammal, comprises
 PT administering to the mammal one or more zona pellucida peptide.

XX Example 1; Col 9; 33pp; English.

PS

XX This invention relates to a novel human anti-zona pellucida protein (ZP1)
CC antibodies in a mammal. The method comprises administering to the mammal
CC one or more zona pellucida peptides shown in the specification. The
CC invention also discloses a method for preventing sperm binding to zona
CC pellucida in a mammal, this may be used as a contraceptive or to create a
CC vaccine. The method of the invention is useful for inducing antibodies to
CC the zona pellucida protein, thus, resulting in a contraceptive effect in
CC multiple mammalian species. The present sequence represents a human zona
CC pellucida immunogenic peptide used to generate the specific antibodies of
CC the invention
XX
SQ Sequence 19 AA;
QY Query Match 34.3%; Score 34; DB 6; Length 19;
Best Local Similarity 53.8%; Pred. No. 3.7e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
1 SSDGLMNNQTOL 13
1 SSGSGIDNYQTOL 13
Db
RESULT 9
ABP83017
ID ABP83017 standard; peptide; 20 AA.
XX
AC ABP83017;
XX
DT 04-MAR-2003 (first entry)
XX
DE G protein-coupled receptor (GPCR) antigenic peptide SEQ ID NO:1690.
XX
KM G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KM G protein-coupled receptor modulator; antibody; immune-related disease;
KM growth-related disease; cell regeneration-related disease; AIDS; cancer;
KM immunological-related cell proliferative disease; autoimmune disease;
KM Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KM osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KM graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KM psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KM mental retardation; epilepsy; aschma; tuberculosis; obesity; nausea;
KM hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
ulcer.
XX
OS Homo sapiens.
XX
PN WO200261087-A2.
XX
PD 08-AUG-2002.
XX
PF 19-DEC-2001; 2001WO-US050107.
XX
PR 19-DEC-2000; 2000US-0257144P.
XX
PA (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
PI Burner GC, Roush CL, Brown JP;
XX
XX WPI; 2003-046718/04.
XX
DR
XX
PT New isolated antigenic peptides e.g., for G protein-coupled receptors
PT (GPCR), useful for diagnosing and designing drugs for treating conditions
PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
PT autoimmune diseases.
XX
PS Claim 1; Fig 2; 523pp; English.
XX
CC The present invention describes antigenic peptides (I) comprising: (a)
CC any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
CC acids. Also described: (1) an assay for the detection of a particular G
CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC and (2) an isolated antibody having high specificity and high affinity or

CC avidity for a particular GPCR. (I) can be used as GPCR modulators and in
CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an
CC antibody against a particular GPCR, and in the production of specific
CC antibodies. The peptides and antibodies are also useful for detecting the
CC presence or absence of corresponding GPCRs. The antigenic peptides for
CC GPCRs and antibodies are useful for diagnosing and designing drugs for
CC treating immune-related diseases, growth-related diseases, cell
CC regeneration-related disease, immunological-related cell proliferative
CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC any other disorder in which GPCRs are involved. The antibodies may be
CC used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42865 encode
CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
CC exemplification of the present invention
XX
SQ Sequence 20 AA;
QY Query Match 34.3%; Score 34; DB 6; Length 20;
Best Local Similarity 66.7%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
1 SSDGLMNNN 9
12 SSDSDMNNN 20
Db
RESULT 10
ADE11136
ID ADE11136 standard; peptide; 18 AA.
XX
AC ADE11136;
XX
DT 29-JAN-2004 (first entry)
XX
DE Antigenic peptide B1-92-109 #SEQ ID 18.
XX
KM Antigenic peptide; omega-oxidation; fatty acid; alkane; alpha;
KM omega-dicarboxylic acid; yeast; antibody; perfume; polymer; adhesive;
KM macrolide antibiotic.
XX
OS Candida tropicalis.
XX
PN WO2003089611-A2.
XX
PD 30-OCT-2003.
XX
PF 18-APR-2003; 2003WO-US012158.
XX
PR 19-APR-2002; 2002US-0374066P.
XX
PR 18-APR-2003; 2003US-00418820.
XX
PA (COGN-) COGNIS CORP.
XX
PI Zhang Y, Wilson CR;
XX
XX WPI; 2004-011692/01.
XX
DR
XX
PT New antigenic peptide, useful for producing antibodies specific for yeast
PT enzymes involved in dicarboxylic acid biosynthesis, for monitoring these
PT enzymes.
XX
PS Claim 1; SEQ ID NO 18; 66pp; English.
XX
CC The invention relates to antigenic peptides having any of seven specific
CC sequences (of 16-25 amino acids, reproduced) or their analogs,
CC derivatives and immunologically active fragments. They are used to raise
CC antibodies (Ab) that bind selectively to enzymes (II) involved in the

CC omega-oxidation of fatty acids or alkanes to alpha,omega-dicarboxylic
CC acids in yeast. Antigenic peptides are used to raise antibodies (Ab)
CC specific for yeast enzymes (II) involved in the omega-oxidation of fatty
CC acids or alkanes to alpha,omega-dicarboxylic acids (III). Antibodies are
CC used to detect/quantify (II), especially for monitoring their induction
CC and/or stability during oxidation to (III), e.g. for manipulating (II)
CC for increased yield of (III), and also to verify the effect of
CC modifications of yeast enzyme encoding genes. Alpha,omega-dicarboxylic
CC acids are useful as intermediates for perfumes, polymers, adhesives and
CC macrolide antibiotics. The current sequence represents an antigenic
CC peptide designed for recombinant antibody generation.

SQ Sequence 18 AA;

Query Match 33.3%; Score 33; DB 8; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GLMNN 8
| | | |
Db 1 GLMNN 5

RESULT 11

ID ABG80887 standard; protein; 20 AA.

AC ABG80887;

DT 29-NOV-2002 (first entry)

DE Human es1 mutant protein E107N.

XX Scaffold protein; C-type lectin-like domain; CTLD; alpha-helix;
KW beta-strand; connecting segment; 14loop region; tetranectin;
KM ligand-binding specificity; human; mouse; rat; chicken; mutant; mutein.

XX Homo sapiens.

OS Synthetic.

PN WO200248189-A2.

PD 20-JUN-2002.

PF 13-DEC-2001; 2001WO-DK000825.

PR 13-DEC-2000; 2000DK-00001872.

PR 28-FEB-2001; 2001US-0272098P.

PA (BORE-) BOREAN PHARMA AS.

PI Etzerodt M, Holter TL, Graversen NH, Thogersen HC;

DR WPI; 2002-643278/69.

XX Protein comprising a variant of model C-type lectin-like domains (CTLD),
PT in which alpha helices, beta-strands, connecting segments are conserved
PT to maintain CTLD scaffold structure, while the loop region is altered.

PS Claim 1; Page 19; 168pp; English.

CC The present invention relates to a new protein with scaffold structure of
CC C-type lectin-like domains (CTLD). The invention comprises a variant of a
CC model CTLD where alpha-helices and beta-strands and connecting segments
CC are conserved such that scaffold structure of C-type lectin-like domains
CC (CTLD) is substantially maintained, while the 14loop region is altered by
CC amino acid substitution, deletion, insertion or their combination. The
CC invention is useful for preparing a library of nucleotide sequences
CC encoding related proteins by randomising part or all of the nucleic acid
CC sequence encoding the loop region of its CTLD. The artificial CTLD
CC protein products are preferable to antibody derivatives as each binding
CC site is a single structurally autonomous protein domain. When used as
CC components of compositions to be used for in vivo diagnostic or

CC therapeutic purposes, artificial CTLD protein products constructed on the
CC basis of human CTLDs are virtually identical to the corresponding natural
CC CTLD protein already present in the body and are therefore less
CC immunogenic to the patient. They also have a smaller size, and thus
CC provide tissue penetration and distribution, as well as shorter half life
CC in circulation. Since murine and human tetranectin are identical in
CC structure, straightforward swapping of polypeptide segments defining
CC ligand-binding specificity between murine and human tetranectin
CC derivatives may be achieved. The present amino acid sequence represents a
CC mutant protein of the invention

SQ Sequence 20 AA;

Query Match 33.3%; Score 33; DB 5; Length 20;
Best Local Similarity 57.1%; Pred. No. 5.5e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 GLMNNQ 10
| | | |
Db 13 GMMNDNR 19

RESULT 12

ID ABO14120 standard; peptide; 20 AA.

AC ABO14120;

DT 21-AUG-2003 (first entry)

DE Novel human secreted protein fragment #21.

XX Human; secreted protein; cytostatic; neuroprotective; hepatotropic;
KW gene therapy; cancer; liver disorder; hepatitis; neural disorder;
KM Alzheimer's disease.

XX Homo sapiens.

PN US2003028003-A1.

PD 06-FEB-2003.

PF 12-OCT-2001; 2001US-00974879.

PR 07-NOV-1997; 97US-0064908P.

PR 07-NOV-1997; 97US-0064908P.

PR 07-NOV-1997; 97US-0064911P.

PR 07-NOV-1997; 97US-0064912P.

PR 07-NOV-1997; 97US-0064983P.

PR 07-NOV-1997; 97US-0064984P.

PR 07-NOV-1997; 97US-0064985P.

PR 07-NOV-1997; 97US-0064987P.

PR 17-NOV-1997; 97US-0066088P.

PR 17-NOV-1997; 97US-0066089P.

PR 17-NOV-1997; 97US-0066090P.

PR 17-NOV-1997; 97US-0066094P.

PR 17-NOV-1997; 97US-0066095P.

PR 17-NOV-1997; 97US-0066100P.

PR 04-NOV-1998; 98WO-US023435.

PR 05-MAY-1999; 99US-00305736.

PR 13-OCT-2000; 2000US-0239893P.

PR 28-MAR-2001; 2001US-00818683.

XX (ROSE/) ROSEN C A.

XX (FENG/) FENG P.

XX (RUBE/) RUBEN S M.

XX (EBNE/) EBNER R.

XX (OLSE/) OLSEN H S.

XX (NIJ/) NI J.

XX (WEI/) WEI Y.

XX (SOP/) SOPPER D R.

XX (MOOR/) MOORE P A.

XX (KYAW/) KYAW H.

PA (LAFLEU/) LAFLEUR D W.
 PA (SHIY/) SHI Y.
 PA (JANAT/) JANAT F.
 PA (ENDR/) ENDRESS G A.
 PA (CART/) CARTER K C.
 PA (BIRSE/) BIRSE C E.
 XX
 PI Rosen CA, Feng P, Ruben SM, Ebner R, Olsen HS, Ni J, Wei Y,
 PI Soppet DR, Moore PA, Kyaw H, Lafleur DW, Shi Y, Janat F,
 PI Endress GA, Carter KC, Birse CE;
 XX MPI, 2003-479549/45.
 XX
 PT New nucleic acid molecule, useful for preparing a medicament for
 PT preventing, treating or ameliorating a medical condition e.g., cancer,
 PT liver disorders such as hepatitis or neural disorders such as Alzheimer's
 PT disease.
 XX
 PS Disclosure: Page 9; 496pp; English.
 XX
 CC The invention describes a new isolated nucleic acid molecule comprising a
 CC sequence having at least 95% identity with a sequence comprising: (a) a
 CC polynucleotide (PN) fragment of a sequence comprising 420-3435 bp, or its
 CC allelic variant; (b) a PN fragment of the cDNA sequence; (c) a PN
 CC sequence encoding a polypeptide, or its fragment, domain, epitope or
 CC species homologue; or (d) a PN that hybridises under stringent conditions
 CC to any one of the sequences of (A)-(C). The nucleic acid is useful for
 CC preparing a medicament for preventing, treating or ameliorating a medical
 CC condition e.g., cancer, liver disorders such as hepatitis or neural
 CC disorders such as Alzheimer's disease. This is the amino acid sequence of
 CC a novel human secreted protein fragment
 XX
 SQ Sequence 20 AA;
 XX
 Query Match 33.3%; Score 33; DB 6; Length 20;
 Best Local Similarity 50.0%; Pred. No. 5.5e+02;
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 5 LMNNNGTQLFLEHS 18
 | | | | | : :
 Db 5 LVNNGTQRLLOEA 18

RESULT 13
 ADG78515
 ID ADG78515 standard; peptide; 20 AA.
 XX
 AC ADG78515;
 XX
 DT 11-MAR-2004 (first entry)
 XX
 DE Human secreted protein #138.
 XX
 KW Human; secreted protein; immune disorder; hematopoietic disorder;
 KW hyperproliferative disorder; infectious disease; inflammatory disorder;
 KW HIV; anaemia; thrombocytopenia; bleeding; stroke; myocardial infarction;
 KW Addison's disease; rheumatoid arthritis; dermatitis; Grave's disease;
 KW multiple sclerosis; glomerulonephritis; diabetes;
 KW graft-versus-host disease; inflammatory bowel disease; cancer;
 KW bacterial infection; viral infection; parasitic infection; osteoporosis;
 KW wound; atherosclerosis; Alzheimer's disease; Parkinson's disease; ulcer;
 KW food additive; food preservative.
 KW
 XX Homo sapiens.
 XX OS
 XX PN US2003211472-A1.
 XX PD 13-NOV-2003.
 XX PF 28-MAR-2001; 2001US-00818683.
 XX PR 05-MAY-1999; 99US-00305736.

PA (FENG/) FENG P.
 PA (ROBE/) ROSEN S M.
 PA (ROSE/) ROSEN C A.
 PA (EBNER/) EBNER R.
 PA (OLSEN/) OLSEN H S.
 PA (NIJ/) NI J.
 PA (WEIY/) WEI Y.
 PA (SOPP/) SOPPET D R.
 PA (MOOR/) MOORE P A.
 PA (KYAW/) KYAW H.
 PA (LAFLEU/) LAFLEUR D W.
 PA (SHIY/) SHI Y.
 PA (JANAT/) JANAT F.
 PA (ENDR/) ENDRESS G A.
 PA (CART/) CARTER K C.
 XX
 PI Peng P, Ruben SM, Rosen CA, Ebner R, Olsen HS, Ni J, Wei Y;
 PI Soppet DR, Moore PA, Kyaw H, Lafleur DW, Shi Y, Janat F;
 PI Endress GA, Carter KC;
 XX MPI; 2004-051575/05.
 XX
 CC New secreted nucleic acid for diagnosing, preventing or treating diseases
 CC associated with aberrant expression or activity of the polypeptide it
 CC encodes, e.g. cancer, human immunodeficiency virus, Parkinson's disease,
 CC or diabetes.
 XX
 PS Disclosure; SEQ ID NO 276; 377pp; English.
 XX
 CC The invention also relates to human secreted proteins and the nucleic
 CC acids encoding them. The proteins and nucleic acids are useful in
 CC diagnosing, preventing, prognosing or treating diseases or disorders
 CC associated with aberrant expression and/or activity of the secreted
 CC proteins, such as immune disorders, hematopoietic disorders,
 CC hyperproliferative disorders, infectious diseases or inflammatory
 CC disorders. In particular, the diseases or disorders are HIV, anaemia,
 CC thrombocytopenia, bleeding, stroke, myocardial infarction, Addison's
 CC disease, rheumatoid arthritis, dermatitis, Grave's disease, multiple
 CC sclerosis, glomerulonephritis, diabetes, graft-versus-host disease,
 CC inflammatory bowel disease, cancer, bacterial infections, viral
 CC infections, parasitic infections, osteoporosis, wounds, atherosclerosis,
 CC Alzheimer's disease, Parkinson's disease or ulcers. The sequences may
 CC also be used as food additives or preservatives, or for modulating
 CC mammalian mental or physical characteristics. The nucleic acids are also
 CC used in chromosome mapping, in forensic biology or as molecular weight
 CC markers. This sequence represents a human secreted protein of the
 CC invention.
 XX
 SQ Sequence 20 AA;
 XX
 Query Match 33.3%; Score 33; DB 8; Length 20;
 Best Local Similarity 50.0%; Pred. No. 5.5e+02;
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 5 LMNNNGTQLFLEHS 18
 | | | | | : :
 Db 5 LVNNGTQRLLOEA 18

RESULT 14
 ADN60818
 ID ADN60818 standard; peptide; 20 AA.
 XX
 AC ADN60818;
 XX
 DT 01-JUL-2004 (first entry)
 XX
 DE Human secreted polypeptide #149.
 XX
 KW Human; secreted polypeptide; autoimmune disease; rheumatoid arthritis;
 KW hyperproliferative disorder; neoplasm; breast; liver;
 KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
 KW cerebral ischaemia; nervous system disorder; Alzheimer's disease;

KW bacterial infection; viral infection; fungal infection; ocular disorder;
 KW corneal infection; wound healing; epithelial cell proliferation;
 KW skin aging; sunburn; organ transplantation; tissue regeneration;
 KW chemotaxis; food additive; food preservative; fat content; vitamin;
 KW mineral.

XX Homo sapiens.

XX US2004038277-A1.

XX 26-FEB-2004.

XX 18-JUL-2003; 2003US-00621401.

XX 07-NOV-1997; 97US-0064900P.

XX 07-NOV-1997; 97US-0064908P.

XX 07-NOV-1997; 97US-0064911P.

XX 07-NOV-1997; 97US-0064912P.

XX 07-NOV-1997; 97US-0064983P.

XX 07-NOV-1997; 97US-0064984P.

XX 07-NOV-1997; 97US-0064985P.

XX 07-NOV-1997; 97US-0064987P.

XX 07-NOV-1997; 97US-0064988P.

XX 17-NOV-1997; 97US-0066089P.

XX 17-NOV-1997; 97US-0066090P.

XX 17-NOV-1997; 97US-0066094P.

XX 17-NOV-1997; 97US-0066095P.

XX 17-NOV-1997; 97US-0066100P.

XX 04-NOV-1998; 98WO-US023435.

XX 05-MAY-1999; 99US-00305736.

XX 13-OCT-2000; 2000US-0239893P.

XX 28-MAR-2001; 2001US-00818683.

XX 12-OCT-2001; 2001US-00974879.

XX (HUMA-) HUMAN GENOME SCT INC.

XX Rosen CA, Peng P, Ruben SM, Ebner R, Olsen H, Ni J, Wei Y;

XX Soppe DR, Moore PA, Kyaw H, Lafleur DM, Shi Y, Janat F;

XX Endress GB, Carter KC, Birse CE;

XX WPI; 2004-203215/19.

XX New nucleic acid molecule encoding one of 125 human secreted proteins for

XX diagnosing, preventing, treating or ameliorating medical conditions and

XX used as food additives or preservatives.

XX Disclosure; SEQ ID NO 288; 531pp; English.

XX The invention relates to isolated nucleic acid molecules encoding human

XX secreted proteins. The polynucleotides and polypeptides are used in

XX diagnosing a pathological condition or susceptibility to a pathological

XX condition. Antibodies to the polypeptides can also be used in alleviating

Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 5,IMNNNOTQLPFEHS 18
 | | | | | : : :
 Db 5 LVYNNQTORLLQEA 18

RESULT 15

AA091293

ID AAR91293 standard; peptide; 12 AA.

XX AAR91293;

XX 14-OCT-1996 (first entry)

XX Anti-idiocytic T-cell modulating peptide.

XX Peptide; VDJ; anti-idiocytic T cell; vaccine; detection; diagnosis;

XX insulin dependent diabetes mellitus; IDDM; assay; proliferation;

XX cytokine.

XX Synthetic.

XX WO9611214-A1.

XX 18-APR-1996.

XX 10-OCT-1995; 95WO-US012686.

XX 07-OCT-1994; 94IL-00111196.

XX (YEDA) YEDA RES & DEV CO LTD.

XX Cohen IR, Elias D;

XX WPI; 1996-209811/21.

XX Novel VDJ peptide and corresponding DNA - used in treatment and

XX prevention of insulin dependent diabetes mellitus.

XX Disclosure; Page 41; 60pp; English.

XX Peptides having a VDJ region where V includes the dipeptide sequence A-S,

XX D preferably has 2-5 amino acids and includes the dipeptide L-G and J

XX includes the tripeptide N-Q-D. may be used as agents for the detection of

XX anti-idiocytic T-cells and in a vaccine against insulin dependent

XX diabetes mellitus (IDDM). The peptides may also be used in the prevention

XX and treatment of IDDM by activating autologous T-cells against the

XX peptides and then re-administering them to the patient. The peptides may

XX also be used in the diagnosis or staging of IDDM or for monitoring the

XX course of treatment of IDDM by assaying T-cells of the subject being

XX tested for proliferation or cytokine production upon in vitro contact

XX with the peptides

XX Sequence 12 AA;

XX Query Match 32.3%; Score 32; DB 2; Length 12;

XX Best Local Similarity 45.5%; Pred. No. 4.5e+02;

XX Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 SDGLMNNQQTQ 12
 : | | | | |
 Db 1 ASLMTNODTQ 11

RESULT 16

AA09338

ID AAM09338 standard; peptide; 18 AA.

XX AAM09338;

XX 16-OCT-2003 (revised)

XX 29-MAY-1997 (first entry)

XX FIV principal immunodominant domain in a HIV-1 context.
 DE
 XX
 XX Feline immunodeficiency virus; principal immunodominant domain; mutant;
 KW lentivirus; envelope protein; neutralising antibody; vaccine.
 XX
 XX Feline immunodeficiency virus.
 OS Human immunodeficiency virus 1.
 OS Chimeric.
 XX
 PN W09630527-A1.
 PD
 PD 03-OCT-1996.
 XX
 XX 26-MAR-1996; 96WO-FR000449.
 XX
 XX 27-MAR-1995; 95FR-00003566.
 XX
 PA (CNRS) CNRS CENT NAT RECH SCT.
 XX
 PI Pancino G, Sonigo P;
 XX
 XX WPI; 1996-455372/45.
 DR
 DR N-PSDB; AAT48025.
 XX
 PT Mutant lenti-virus Env protein principal immuno-dominant domains - can
 PT produce neutralising antibodies to protect against e.g. FIV or HIV, but
 PT do not produce facilitating antibodies.
 XX
 XX Claim 3; Page 19; 38pp; French.
 XX
 XX The present sequence is a specifically claimed example of a novel peptide
 CC fragment which is a mutant of the wild-type (w.t.) principal
 CC immunodominant domain (PID) of human immunodeficiency virus such that an
 CC Env protein containing the fragment retains the ability to produce
 CC neutralising antibodies (Abs) but does not produce facilitating or
 CC deleterious Abs against the w.t. PID. The mutated Env protein can be used
 CC for producing vaccines against HIV. In addition to vaccination, the
 CC peptide fragment is useful as a diagnostic reagent, specifically to
 CC distinguish between Abs produced in response to vaccination and those
 CC resulting from viral infection. Also, the results of anti-HIV vaccination
 CC can be monitored using the peptide. (Updated on 16-OCT-2003 to
 CC standardise OS field)
 CC
 XX
 SQ Sequence 18 AA;
 QY
 QY 5 LMNNNQTLF 14
 : | | | | |
 1 IWGCNQNOFF 10
 Db
 RESULT 17
 AA07984
 ID AA07984 standard; peptide; 19 AA.
 XX
 AC AA07984;
 XX
 XX 16-OCT-2003 (revised)
 DT 03-FEB-1997 (first entry)
 XX
 DE gp41 peptide MN-13.
 XX
 XX HIV; gp120; HIV-IIIB strain; HIV-1 transmission; foetal transmission;
 KW neutralising antibody; passive immunisation; anti-idiotypic antibody;
 KW gp41; vaccine; active immunotherapy.
 XX
 OS Human immunodeficiency virus 1.
 XX
 PN US5556744-A.

XX
 PD 17-SEP-1996.
 XX
 XX 24-MAR-1994; 94US-00218025.
 PE
 XX 29-MAY-1992; 92US-00891451.
 FR
 XX (TYPE-) UNIV PENNSYLVANIA.
 PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
 XX
 PI Williams WV, Weiner DB, Ugen KE;
 XX
 XX WPI; 1996-432980/43.
 DR
 XX
 XX Determining the likelihood of maternal transmission of HIV-1 to foetus -
 PT by measuring maternal reactivity with specific gp120 and gp41 derived
 PT peptide(s), also used for diagnosing HIV in infants.
 XX
 XX Example 7; Col 101-102; 63pp; English.
 PS
 PS AA07980-W07987 represent HIV gp41 peptides that can be used in the
 CC method of the invention. The method of the invention is for determining
 CC whether or not a mother will transmit HIV-1 to a foetus. The method
 CC comprises incubating a sample from the HIV-infected mother, with a
 CC collection of HIV peptides. The HIV peptides included at least one of
 CC these sequences, and at least one HIV gp120 derived peptide (see AA07909
 CC -W07917). The number of peptides that react with the sample is
 CC determined, and this number is compared with a standard that shows
 CC pattern reactivity for a patient of transmission status. A non-
 CC transmissive HIV sample is indicated if the test sample reacts with twice
 CC as many peptides as the standard. The method detects the presence of
 CC neutralising antibodies that protect against mother to infant
 CC transmission of HIV. These sequences can also be used in vaccines to
 CC protect against transmission. Antibodies against these sequences can be
 CC used for passive immunisation, and to generate anti-idiotypic antibodies
 CC for use in vaccines or active immunotherapy. (Updated on 16-OCT-2003 to
 CC standardise OS field)
 CC
 XX
 SQ Sequence 19 AA;
 QY
 QY 1 SSDGIANN 8
 : | | | | |
 1 SDDIANN 8
 Db
 RESULT 18
 AA64600
 ID AA64600 standard; peptide; 15 AA.
 XX
 AC AA64600;
 XX
 XX 31-OCT-2001 (first entry)
 DT
 XX
 DE Human excitatory amino acids transporter 17 peptide.
 XX
 XX Human; excitatory amino acid transporter 17; cytostatic; virucidal;
 KW immunomodulatory; antiinflammatory; haemostatic; malignant tumour; HIV;
 KW human immunodeficiency virus; infection; immunological disease;
 KW substance metabolic disorder; neural mental illness;
 KW embryonic development disorder.
 XX
 OS Homo sapiens.
 XX
 XX W020016726-A1.
 PN
 XX 13-SEP-2001.
 PD
 XX 26-FEB-2001; 2001WO-CN000172.
 PF
 XX

PR 10-MAR-2000; 2000CN-0011957.
XX
XX (SHAN-) SHANGHAI BIOWINDOW GENE DEV INC.
XX
XX Mao Y, Xie Y;
XX
XX WPI; 2001-557864/62.
XX
XX Human excitatory amino-acid transporter 17 and encoded polynucleotide,
PT applicable in diagnosis and treatment of e.g. malignant tumor, hemopathy,
PT HIV infection, immunological diseases and inflammations.
XX
XX Example 5; Page 19; 36pp; Chinese.
XX
XX The invention relates to the human excitatory amino-acid transporter 17
CC with cytotoxic, virucidal, immunomodulatory, antiinflammatory and
CC haemostatic activity. The protein and encoding polynucleotide are
CC applicable in the diagnosis and treatment of malignant tumours,
CC haemopathy, HIV infection, immunological diseases, various inflammations,
CC substance metabolic disorder, neural mental illness, embryonic
CC development disorder and growth and development disturbance illness. The
CC present sequence is that of a human excitatory amino-acid transporter 17
CC peptide
XX
XX Sequence 15 AA;
SQ
Query Match 31.3%; Score 31; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 SDGLW 6
DB 8 SDGLW 12
RESULT 19
ADJ38603
ID ADJ38603 standard; peptide; 15 AA.
XX
XX ADJ38603;
AC
XX
XX 06-MAY-2004 (first entry)
DT
XX
XX HSV-4 BFLF2 protein late domain motif mutant peptide #2.
DE
XX
XX Virucide; HSV infection; antiviral; late domain motif; mutein; mutant.
XX
XX Human herpesvirus 4.
OS
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FT Misc-difference 6 /note= "Wild-type residue replaced with Arg"
FT
XX
XX WO2004009027-A2.
XX
XX 29-JAN-2004.
PD
XX
XX 21-JUL-2003; 2003WO-US022828.
XX
XX 19-JUL-2002; 2002US-0397265P.
XX
XX 19-JUL-2002; 2002US-0397477P.
XX
XX 19-JUL-2002; 2002US-0397479P.
XX
XX 03-MAR-2003; 2003US-0451903P.
XX
XX (MYRI-) MYRIAD GENETICS INC.
XX
XX Morham S, Zavitz K, Hobden A;
XX
XX WPI; 2004-123282/12.
XX
XX Use of cells displaying herpes simplex virus (HSV) altered budding
PT phenotype for the manufacture of a medicament for treating HSV infection.
PT

XX
XX Example 14; Page 68; 74pp; English.
PS
XX
XX The present invention relates to cells displaying herpes simplex virus
CC (HSV) altered budding phenotype which are useful for the manufacture of a
CC medicament for treating HSV infection. The medicament further comprises
CC an adjuvant capable of stimulating cellular immune response e.g. IL-2, IL
CC -4, IL-12, IL-18 or gamma-interferon. The cells are human cells and
CC contain a nucleic acid encoding a mutant HSV protein or a nucleic acid
CC encoding the polypeptide sufficient for virus-like particle assembly but
CC devoid of late-domain motifs. The nucleic acid is within an HSV genome.
CC The genome is devoid of late domain motifs capable of effecting viral
CC budding. The composition also comprises a compound capable of interfering
CC with the protein-protein interaction between a host cell protein capable
CC of binding a late domain motif and a HSV protein containing a late domain
CC motif. The composition further comprises another HSV protein or its
CC immunogenic fragment, and/or a nucleic acid encoding the other HSV
CC protein or the immunogenic fragment. The present sequence is a mutant HSV
CC peptide, derived from the wild-type peptide ADJ38601, used to illustrate
CC the invention.
XX
XX Sequence 15 AA;
SQ
Query Match 31.3%; Score 31; DB 8; Length 15;
Best Local Similarity 66.7%; Pred. No. 8.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 9 NOTQFLFH 17
DB 5 NRTQFLFH 13
RESULT 20
ADL25941
ID ADL25941 standard; peptide; 15 AA.
XX
XX ADL25941;
AC
XX
XX 17-JUN-2004 (first entry)
DT
XX
XX Synthetic peptide A128 derived from a conserved region of HCV.
DE
XX
XX HCV; hepatitis C virus; virucide; vaccine; MHC; HLA;
XX major histocompatibility complex; human leukocyte antigen.
XX
XX Synthetic.
OS
XX
XX WO2004024182-A2.
XX
XX 25-MAR-2004.
XX
XX 27-AUG-2003; 2003WO-EP009482.
XX
XX 13-SEP-2002; 2002AT-00001376.
XX
XX 27-FEB-2003; 2003WO-EP002005.
XX
XX 11-JUL-2003; 2003EP-00450171.
XX
XX (INTE-) INTERCELL AG.
XX
XX Buschle M, Habel A, Klade C, Matner F, Ocava O, Vycvyska O;
XX
XX Zauner W, Zinke S, Kiriappos H;
XX
XX WPI; 2004-269899/25.
XX
XX Isolating Hepatitis C virus peptides (Hps) which have a binding capacity
PT to a MHC/HLA molecule or a complex comprising the HCV-peptide and the
PT molecule by separating the complex from the HCV-peptides which do not
PT bind to the molecule.
XX
XX Example 1; Page 30; 73pp; English.
XX
XX The invention relates to a novel method for isolating Hepatitis C Virus
CC (HCV) peptides (Hps). The method of the invention has virucide activity,
CC

CC and may be useful in producing a vaccine. The method is useful for
CC isolating Hepatitis C Virus peptides (Hps) which have a binding capacity
CC to a MHC/HLA molecule or a complex comprising the HCV-peptide and the
CC MHC/HLA molecule for preparing a vaccine against HCV infection. The T
CC cells, a T cell clone or a T cell population or preparation is useful for
CC identifying heteroclitic epitopes or for preparing a composition for
CC treating HCV infection. The present sequence represents a synthetic
CC peptide derived from a conserved region of HCV.

XX Sequence 15 AA;

Query Match 31.3%; Score 31; DB 8; Length 15;
Best Local Similarity 38.5%; Pred. No. 8.1e+02;

Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SSDGLMNNQQL 13
Db 2 NTNGSMHINRTAL 14

RESULT 21

ADL25931
ID ADL25931 standard; peptide; 15 AA.

AC ADL25931;

DT 17-JUN-2004 (first entry)

XX Synthetic peptide A118 derived from a conserved region of HCV.

XX HCV; hepatitis C virus; virucide; vaccine; MHC; HLA;

KW major histocompatibility complex; human leukocyte antigen.

XX OS Synthetic.

XX WO2004024182-A2.

PD 25-MAR-2004.

PF 27-AUG-2003; 2003WO-EP009482.

XX 13-SEP-2002; 2002AT-00001376.

PR 27-FEB-2003; 2003WO-EP002005.

PR 11-JUL-2003; 2003EP-00450171.

XX (INTE-) INTERCELL AG.

PA Bueschle M, Habel A, Klade C, Mattner F, Otava O, Vytvytska O;

PI Zauner W, Zinke S, Kirilappos H;

XX WPI; 2004-269899/25.

XX Isolating Hepatitis C Virus peptides (Hps) which have a binding capacity
PT to a MHC/HLA molecule or a complex comprising the HCV-peptide and the
PT molecule by separating the complex from the HCV-peptides which do not
PT bind to the molecule.

PS Example 1; Page 30; 73pp; English.

XX The invention relates to a novel method for isolating Hepatitis C Virus
CC (HCV) peptides (Hps). The method of the invention has virucide activity,
CC and may be useful in producing a vaccine. The method is useful for
CC isolating Hepatitis C Virus peptides (Hps) which have a binding capacity
CC to a MHC/HLA molecule or a complex comprising the HCV-peptide and the
CC MHC/HLA molecule for preparing a vaccine against HCV infection. The T
CC cells, a T cell clone or a T cell population or preparation is useful for
CC identifying heteroclitic epitopes or for preparing a composition for
CC treating HCV infection. The present sequence represents a synthetic
CC peptide derived from a conserved region of HCV.

XX Sequence 15 AA;

Query Match 31.3%; Score 31; DB 8; Length 15;

Best Local Similarity 38.5%; Pred. No. 8.1e+02;
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SSDGLMNNQQL 13
Db 3 NTNGSMHINRTAL 15

RESULT 22

ADL25931
ID ADL25931 standard; peptide; 15 AA.

AC ADL25931;

DT 17-JUN-2004 (first entry)

XX Synthetic peptide A120 derived from a conserved region of HCV.

XX HCV; hepatitis C virus; virucide; vaccine; MHC; HLA;

KW major histocompatibility complex; human leukocyte antigen.

XX OS Synthetic.

XX WO2004024182-A2.

PD 25-MAR-2004.

PF 27-AUG-2003; 2003WO-EP009482.

XX 13-SEP-2002; 2002AT-00001376.

PR 27-FEB-2003; 2003WO-EP002005.

PR 11-JUL-2003; 2003EP-00450171.

XX (INTE-) INTERCELL AG.

PA Bueschle M, Habel A, Klade C, Mattner F, Otava O, Vytvytska O;

PI Zauner W, Zinke S, Kirilappos H;

XX WPI; 2004-269899/25.

XX Isolating Hepatitis C Virus peptides (Hps) which have a binding capacity
PT to a MHC/HLA molecule or a complex comprising the HCV-peptide and the
PT molecule by separating the complex from the HCV-peptides which do not
PT bind to the molecule.

PS Example 1; Page 30; 73pp; English.

XX The invention relates to a novel method for isolating Hepatitis C Virus
CC (HCV) peptides (Hps). The method of the invention has virucide activity,
CC and may be useful in producing a vaccine. The method is useful for
CC isolating Hepatitis C Virus peptides (Hps) which have a binding capacity
CC to a MHC/HLA molecule or a complex comprising the HCV-peptide and the
CC MHC/HLA molecule for preparing a vaccine against HCV infection. The T
CC cells, a T cell clone or a T cell population or preparation is useful for
CC identifying heteroclitic epitopes or for preparing a composition for
CC treating HCV infection. The present sequence represents a synthetic
CC peptide derived from a conserved region of HCV.

XX Sequence 15 AA;

Query Match 31.3%; Score 31; DB 8; Length 15;
Best Local Similarity 38.5%; Pred. No. 8.1e+02;
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SSDGLMNNQQL 13
Db 1 NTNGSMHINRTAL 13

RESULT 23

ADL25931
ID ADL25931 standard; peptide; 15 AA.

XX

AC ADL25932;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE Synthetic peptide A119 derived from a conserved region of HCV.
 XX
 KW HCV; hepatitis C virus; vircicide; vaccine; MHC; HLA;
 XX major histocompatibility complex; human leukocyte antigen.
 XX
 OS Synthetic.
 XX
 PN MO2004024182-A2.
 XX
 PD 25-MAR-2004.
 XX
 DR 27-AUG-2003; 2003WO-EP009482.
 XX
 PF 13-SEP-2002; 2002AT-00001376.
 XX
 PR 27-FEB-2003; 2003WO-EP002005.
 XX
 PR 11-JUL-2003; 2003EP-00450171.
 XX
 PA (INTE-) INTERCELL AG.
 XX
 PI Bueschle M, Habel A, Klade C, Mattner F, Otava O, Vycvytska O;
 XX Zauner W, Zinke S, Kiriappos H;
 XX
 DR WPI; 2004-269899/25.
 XX
 XX Isolating Hepatitis C Virus peptides (HPs) which have a binding capacity
 PT to a MHC/HLA molecule or a complex comprising the HCV-peptide and the
 PT molecule by separating the complex from the HCV-peptides which do not
 PT bind to the molecule.
 XX
 XX Example 1; Page 30; 73pp; English.
 XX
 CC The invention relates to a novel method for isolating Hepatitis C Virus
 CC (HCV) peptides (HPs). The method of the invention has vircicide activity,
 CC and may be useful in producing a vaccine. The method is useful for
 CC isolating Hepatitis C Virus peptides (HPs) which have a binding capacity
 CC to a MHC/HLA molecule or a complex comprising the HCV-peptide and the
 CC MHC/HLA molecule for preparing a vaccine against HCV infection. The T
 CC cells, a T cell clone or a T cell population or preparation is useful for
 CC identifying heteroclitic epitopes or for preparing a composition for
 CC treating HCV infection. The present sequence represents a synthetic
 CC peptide derived from a conserved region of HCV.
 CC
 XX
 SQ Sequence 15 AA;
 XX
 QY Query Match 31.3%; Score 31; DB 8; Length 15;
 DB Best Local Similarity 38.5%; Pred. No. 8.1e+02;
 Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
 1 SSDGLMNNQOTOL 13
 2 NTNGSMHNRRL 14
 RESULT 24
 ADL25940
 ID ADL25940 standard; peptide; 15 AA.
 XX
 AC ADL25940;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE Synthetic peptide A127 derived from a conserved region of HCV.
 XX
 KW HCV; hepatitis C virus; vircicide; vaccine; MHC; HLA;
 XX major histocompatibility complex; human leukocyte antigen.
 XX
 OS Synthetic.
 XX
 PN MO2004024182-A2.
 XX
 DR

XX
 PD 25-MAR-2004.
 XX
 PF 27-AUG-2003; 2003WO-EP009482.
 XX
 PR 13-SEP-2002; 2002AT-00001376.
 XX
 PR 27-FEB-2003; 2003WO-EP002005.
 XX
 PR 11-JUL-2003; 2003EP-00450171.
 XX
 PA (INTE-) INTERCELL AG.
 XX
 PI Bueschle M, Habel A, Klade C, Mattner F, Otava O, Vycvytska O;
 XX Zauner W, Zinke S, Kiriappos H;
 XX
 DR WPI; 2004-269899/25.
 XX
 XX Isolating Hepatitis C Virus peptides (HPs) which have a binding capacity
 PT to a MHC/HLA molecule or a complex comprising the HCV-peptide and the
 PT molecule by separating the complex from the HCV-peptides which do not
 PT bind to the molecule.
 XX
 XX Example 1; Page 30; 73pp; English.
 XX
 CC The invention relates to a novel method for isolating Hepatitis C Virus
 CC (HCV) peptides (HPs). The method of the invention has vircicide activity,
 CC and may be useful in producing a vaccine. The method is useful for
 CC isolating Hepatitis C Virus peptides (HPs) which have a binding capacity
 CC to a MHC/HLA molecule or a complex comprising the HCV-peptide and the
 CC MHC/HLA molecule for preparing a vaccine against HCV infection. The T
 CC cells, a T cell clone or a T cell population or preparation is useful for
 CC identifying heteroclitic epitopes or for preparing a composition for
 CC treating HCV infection. The present sequence represents a synthetic
 CC peptide derived from a conserved region of HCV.
 CC
 XX
 SQ Sequence 15 AA;
 XX
 QY Query Match 31.3%; Score 31; DB 8; Length 15;
 DB Best Local Similarity 38.5%; Pred. No. 8.1e+02;
 Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
 1 SSDGLMNNQOTOL 13
 3 NTNGSMHNRRL 15
 RESULT 25
 AAMS0838
 ID AAMS0838 standard; peptide; 17 AA.
 XX
 AC AAMS0838;
 XX
 DT 07-AUG-2003 (revised)
 XX
 DT 01-MAY-2002 (first entry)
 XX
 DE HCV E2 region peptide MDL-16 (aa412-429).
 XX
 KW HCV; MDL-16; infection; diagnosis; vaccine.
 XX
 OS Hepatitis C virus.
 XX
 PN MO200204484-A2.
 XX
 PD 17-JAN-2002.
 XX
 PF 06-JUL-2001; 2001WO-CA000988.
 XX
 PR 07-JUL-2000; 2000CA-02311022.
 XX
 PA (MEDM-) MEDMIRA INC.
 XX
 PI Chan HKW, Theolis R;
 XX
 DR WPI; 2002-179696/23.

XX New hepatitis C virus (HCV) peptides and mosaic antigen composition
PT comprising the peptides, useful as immunoreagents for detecting HCV
PT antibodies infection, or as an immunogen for stimulating production of
PT antibodies against HCV.
XX
PS Example 4; Page 33; 59pp; English.
XX
CC The present sequence is that of peptide MDL-16 comprising amino acid
CC residues 412-429 of the hepatitis C virus (HCV) polypeptide E2 region. A
CC panel of HCV polypeptide-derived peptides was tested for immunoreactivity
CC with a commercially available human serum panel. All of the
CC immunoreactivity was concentrated within the core region between amino
CC acids 1 to 80. MDL-16 exhibited low and inconsistent immunogenicity. The
CC invention provides a highly immunoreactive mosaic antigen composition
CC (MAC) comprising 2 or more HCV immunoreactive peptides (see AAM50816-27),
CC each of which is immobilized to a carrier. The unique combination of HCV
CC core peptides in the MAC provides higher specificity and sensitivity for
CC detection of human antibodies specifically to HCV in rapid HCV diagnostic
CC applications. An in vitro diagnostic method for detecting anti-HCV
CC antibodies in a test sample and a diagnostic test kit are provided which
CC use the MAC as an immunoreagent. The HCV MAC may also potentially be used
CC as an immunogen in vaccine compositions. (Updated on 07-AUG-2003 to
CC correct OS field.)
CC
XX Sequence 17 AA;
SQ
Query Match 31.3%; Score 31; DB 5; Length 17;
Best Local Similarity 38.5%; Pred. No. 9.3e+02;
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
QY 1 SSDGIWNNQOTQL 13
Db 4 NTNGSMHINRTAL 16
::|::|::|
X
RESULT 26
ABW01700
ID AEW01700 standard; protein; 19 AA.
XX
XX ABW01700;
AC
XX 12-FEB-2004 (first entry)
DT
XX
XX Toxoplasma gondii TATI-1 transacting factor activation domain.
DE
XX
XX Cellular vaccine; tetracycline repressor; TetR; transacting factor;
KM protozoal infection; gene therapy; TATI.
XX
XX Toxoplasma gondii.
OS
XX US2003185851-A1.
PN
XX 02-OCT-2003.
PD
XX 20-MAR-2002; 2002US-00102143.
PF
XX 20-MAR-2002; 2002US-00102143.
PR
XX 20-MAR-2002; 2002US-00102143.
PS
XX (SOLD/) SOLDATI D.
PA (MEIS/) MEISSNER M.
XX
XX Soldati D, Meissner M;
PI
XX WPI: 2003-844144/78.
DR N-PSDB; AAD63241.
XX
XX Nucleic acid and cellular vaccines useful for treating protozoal
PT infections caused by e.g. Toxoplasma gondii and Plasmodium falciparum.
XX
XX Disclosure; Fig 6A; Opp; English.
XX
XX The present invention relates to nucleic acid and cellular vaccines which

CC comprise tetracycline repressor (TetR) operatively linked to a
CC transacting factor of Toxoplasma gondii. Nucleic acid constructs of the
CC invention are used in medicine as vaccines for the treatment and
CC prevention of protozoal infections caused by Plasmodium falciparum, P.
CC vivax, P. berghei, P. yoelii, P. knowlesi, Entamoeba histolytica,
CC Trypanosoma brucei, T. gondii and/or Giardia lamblia. The invention is
CC useful in gene therapy. The present sequence is T. gondii TATI
CC transacting factor activation domain. This sequence is used in the
CC invention
XX
SQ Sequence 19 AA;
QY Query Match 31.3%; Score 31; DB 7; Length 19;
Best Local Similarity 60.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Db 6 WNNQOTQL 15
8 WNNQPYLGL 17
X
RESULT 27
AD059249
ID AD059249 standard; peptide; 19 AA.
XX
XX AD059249;
AC
XX 15-JUN-2004 (first entry)
DT
XX
XX Toxoplasma gondii TATI-3 activating domain peptide.
DE
XX
XX gene activation; tetracycline-controlled transactivator;
KM tetracycline repressor; TetR; TATI-1; TATI-3; vaccine;
KM parasite infection; protozoacide.
XX
XX Toxoplasma gondii.
OS
XX GB2388112-A.
PN
XX 05-NOV-2003.
PD
XX 20-MAR-2002; 2002GB-00006595.
PF
XX 20-MAR-2002; 2002GB-00006595.
PR
XX 20-MAR-2002; 2002GB-00006595.
PS
XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
PA
XX
XX Soldati D, Meissner M;
PI
XX WPI: 2003-880573/82.
DR N-PSDB; AD059248.
XX
XX New nucleic acid construct comprising tetracycline repressor operatively
PT linked to a transacting factor of Toxoplasma gondii, useful in producing
PT vaccines to treat or prevent infections caused by Apicomplexan parasites.
XX
XX Disclosure; Fig 6b; 76pp; English.
XX
XX This invention relates to a novel control system for regulating gene
CC activation in Toxoplasma gondii (T. gondii) using tetracycline-controlled
CC transactivators. Specifically, it refers to a nucleic acid construct
CC comprising the tetracycline repressor (TetR) operatively linked to a T.
CC gondii transacting factor (TATI-1 or TATI-3) and an activating domain.
CC The present invention describes this construct as useful for preparing a
CC vaccine for the treatment or prophylaxis of an infection caused by a
CC protozoan parasite selected from, for example, T. gondii, Plasmodium
CC falciparum, Trypanosoma brucei, Entamoeba histolytica or Giardia lamblia.
CC Furthermore it can be used to create attenuated strains of the parasites
CC that can act as vaccines to protect against Apicomplexan parasite
CC infection, and also identify essential genes that could provide novel
CC drug targets. As such, these compositions can be referred to as
CC protozoacides. This peptide sequence is the T. gondii TATI-3 activating
CC domain peptide, given in an exemplification of the invention.

XX SQ Sequence 19 AA; 31.3%; Score 31; DB 7; Length 19;
Query Match Best Local Similarity 60.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 WNNNQTLFL 15
|||
8 WNNQPYLGL 17

DB

RESULT 28
ADA26471
ID ADA26471 standard; peptide: 20 AA.
AC ADA26471;
XX
DT 20-NOV-2003 (first entry)
XX
DE Isomaltose manufacture method enzyme-derived peptide #19.
XX
KW isomaltose; alpha-isomaltosyl gluco-sugar synthase;
KW alpha-isomaltosyl transferase; alpha-1; 4-glucosyl bond;
KW alpha-isomaltosyl glucosugar; alpha-1; 6-glucosyl bond; isomaltitol;
KW health food; beverage; fodder; cosmetic; pharmaceutical; humectant;
KW osmotic pressure regulators; sugar crystallization inhibitor;
KW starch aging inhibitor.
XX
OS Sporosarcina globispora.
XX
PN W02003033717-A1.
XX
XX 24-APR-2003.
PD
PF 18-OCT-2002; 2002WO-JP010846.
XX
PR 18-OCT-2001; 2001JP-00321182.
PR 30-AUG-2002; 2002JP-00252609.
XX
PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI Kubota M, Nishimoto T, Sonoda T, Fukuda S, Miyake T;
XX
DR WPI; 2003-430348/40.
XX
PT Manufacture of isomaltose in two-step process from specified sugar types
PT using specified enzymes, for manufacture of isomaltitol and use in
PT foodstuffs, fodder, pharmaceuticals and cosmetics.
XX
PS Disclosure; Page 205; 260pp; Japanese.
XX
XX The invention relates to the manufacture of isomaltose by making one or
CC more alpha-isomaltosyl gluco-sugar synthase from *Bacillus globisporus* N75
CC (FERM BP-7591), *Arthrobacter globiformis* A19 (FERM BP-7590) and
CC *Arthrobacter ramosus* S1 (FERM BP-7592), in the presence or absence of
CC alpha-isomaltosyl transferase from *Bacillus globisporus* N75 (FERM BP-
CC 7591) and/or *Arthrobacter globiformis* A19 (FERM BP-7590) act on a sugar
CC of glucose polymerization degree of two or more having alpha-1,4-glucosyl
CC bonds as the bonding of the non-reducing terminal to give a alpha-
CC isomaltosylglucosugar with glucose polymerization degree of three or more
CC having a alpha-1,6-glucosyl bond as a non-reducing terminal and alpha-1,4
CC -glucosyl bonds as the other terminals, and/or cyclo(-6)-alpha-D-
CC glucopyranosyl-(1-3)-alpha-D-glucopyranosyl-(1-6)-alpha-D-glucopyranosyl-
CC (1-3)-alpha-D-glucopyranosyl-(1-), acting on this with isomaltose release
CC enzyme and collecting the isomaltose. The isomaltose and isomaltitol are
CC used in health foods and beverages, fodder and feeds, cosmetics and
CC pharmaceuticals, and luxury articles, as humectant, osmotic pressure
CC regulators, low sweeteners, sugar crystallization inhibitors, and starch
CC aging inhibitors. This sequence represents a peptide derived from one of
CC the above proteins.
XX
SQ Sequence 20 AA;

XX SQ Sequence 19 AA; 31.3%; Score 31; DB 6; Length 20;
Query Match Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 DGLMNN 8
|||
4 DGVMHN 9

DB

RESULT 29
ADH80940
ID ADH80940 standard; peptide: 20 AA.
XX
AC ADH80940;
XX
DT 15-APR-2004 (first entry)
XX
DE HIV hypervariable epitope construct (HEV) #5.
XX
KW antimicrobial; vaccine; immunogenic peptide mixture; immunogenic epitope;
KW pathogen; reactive immunity; human immunodeficiency virus; HIV;
KW hypervariable epitope construct.
XX
OS Human immunodeficiency virus 1.
XX
FH Key Location/Qualifiers
FT Misc-difference 11 /note= "Thr, Ser"
FT Misc-difference 13 /note= "Ile, Thr"
XX
XX US2002183484-A1.
XX
XX 05-DEC-2002.
PD
PF 08-FEB-2002; 2002US-00072084.
XX
PR 09-OCT-1998; 98US-0103642P.
PR 08-OCT-1999; 99US-00414484.
XX
XX (TORR/) TORRES J V.
XX
XX Torres JV;
XX
XX WPI; 2003-328635/31.
XX
PT Preparation of immunogenic peptide mixtures for e.g. producing vaccines,
PT by obtaining immunogenic epitope sequences having common residue regions
PT and variable residues, and determining frequency of amino acids at the
PT variable residues.
XX
XX Example 1; Fig 3; 30pp; English.
XX
XX The invention describes the preparation of an immunogenic peptide
CC mixture. The mixture is prepared by obtaining immunogenic epitope
CC sequences of a pathogen having a common residue region and at least one
CC variable residue, determining the frequency with which different amino
CC acids are found at the variable residue and synthesizing a peptide
CC mixture comprising up to 100 different peptides using the amino acids.
CC The peptide mixture is useful in the production of vaccines, therapeutic
CC agents and diagnostic kits against pathogenic organisms, e.g. viruses and
CC parasites. The inventive method produces less complex immunogenic peptide
CC formulations than those described in the prior art, while retaining
CC optimum immunogenicity. The resulting immunogenic peptide mixture evokes
CC broadly reactive immunity. This is the amino acid sequence of a human
CC immunodeficiency virus type 1 hypervariable epitope construct used in the
CC immunogenic peptide mixture of the invention.
XX
SQ Sequence 20 AA;

Query Match 31.3%; Score 31; DB 7; Length 20;
Best Local Similarity 50.0%; Pred. No. 1.1e+03;

Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Qy 3 DGLMNNQTLQ 14
|||:|
Db 3 DGNMNNEXEF 14

RESULT 30
ADM29206
ID ADM29206 standard; peptide; 20 AA.
XX
XX
XX

ADM29206;

01-JUN-2004 (first entry)

Hepatitis C virus cellular immune response related peptide.

XX DNA vaccine; vaccine; antigen gene; antigen; hepatitis C virus; HCV;
KW recombinant adenovirus vaccine; adenovirus; HCV infection;
KM hepatitis C virus infection; infection; virucide.
XX
XX

Hepatitis C virus.
OS Synthetic.
XX

WO2004028563-A1.

08-APR-2004.

24-SEP-2003; 2003WO-KR001951.

XX 27-SEP-2002; 2002KR-00058712.
PR 06-NOV-2002; 2002KR-00068496.

XX (GENE-) GENEXINE INC.
PA (POST-) POSTECH FOUND.
PA (DONG-) DONG-A PHARM CO LTD.
PA (DAEW-) DAEWONG CO LTD.
PA (POSC-) POSCO.

Sung YC, Youn J, Yang S, Park S, Lee CG;

WPI; 2004-305120/28.

XX New DNA vaccine comprising plasmid containing 2-6 kb of the total antigen
PT gene of hepatitis C virus (HCV), useful in treating or preventing HCV
infection.
XX

PS Disclosure: Fig 20b; 165pp; English.

XX The present invention describes a DNA vaccine which comprises a plasmid
CC containing 2-6 kb of the total antigen gene of hepatitis C virus (HCV).
CC Also described: (1) a recombinant adenovirus vaccine including an
CC adenovirus containing 2-6 kb of total antigen gene of HCV; (2) a vaccine
CC administering method; and (3) a method for treating or preventing HCV
CC infection. The DNA vaccine has virucide activity. The DNA vaccine is
CC useful in treating or preventing HCV infection. The present sequence
CC represents a peptide from the HCV peptide pool used for the investigation
CC of cellular immune response in the exemplification of the present
CC invention.
XX

Sequence 20 AA;

Query Match 31.3%; Score 31; DB 8; Length 20;
Best Local Similarity 38.5%; Pred. No. 1.1e+03;
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SSDGLMNNQTLQ 13
::||:|
Db 2 NNNGSMHINRTAL 14

RESULT 31
ABO57864

ID ABO57864 standard; protein; 20 AA.

XX ABO57864;

XX 29-JUN-2004 (first entry)

XX Human genome derived single exon protein #4098.

XX Human; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX

OS Homo sapiens.

PN US2003194704-A1.

16-OCT-2003.

03-APR-2002; 2002US-00029386.

03-APR-2002; 2002US-00029386.

PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.

Penn SG, Rank DR, Hanzel DK;

WPI; 2004-119264/12.

XX New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX

PS Claim 45; SEQ ID NO 31498; 80pp; English.

XX The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridizes under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterizing
CC alternative splicing events, in detecting and characterizing gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in printing the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe protein of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docid=20030194704
XX
SQ Sequence 20 AA;

Query Match 31.3%; Score 31; DB 8; Length 20;
Best Local Similarity 60.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 NNNQTQFLFE 16
| | | | | : |
| | | | | : |
DB 6 NYNQQLIME 15

RESULT 32
ADE78906
ID ADE78906 standard; peptide; 10 AA.
XX
AC ADE78906;
XX
DT 29-JAN-2004 (first entry)
XX
DE Interleukin 13 receptor subunit alpha 2 immunogenic peptide.

XX
KM immune response; interleukin 13 receptor subunit alpha 2; IL-13Ralph2;
KM anti-cancer; vaccine; human; immunogenic.
XX
OS Synthetic.
OS Homo sapiens.
PN WO2003092717-A1.
XX
PD 13-NOV-2003.
XX
PF 22-MAR-2002; 2002WO-US008983.
XX
PR 22-MAR-2002; 2002WO-US008983.
XX
PA (PENN-) PENN STATE RES FOUND.
PI Debincki W, Christensen N, Mintz A;
XX
DR WPI; 2003-903570/82.

XX
PT Stimulating an immune response against IL-13Ralph2 in a subject having
PT or at risk for developing a disease, e.g. cancer, by formulating an anti-
PT cancer vaccine comprising an agent that can stimulate the immune response
PT against.
XX
XX
XX Example 3; Page 43; 67p; English.

XX
PS The present invention describes a method for stimulating an immune
CC response against interleukin 13 receptor subunit alpha 2 (IL-13Ralph2)
CC in a subject having or at risk for developing a disease having cells
CC expressing IL-13Ralph2. The method comprises formulating an anti-cancer
CC vaccine comprising an agent that can stimulate an immune response against
CC IL-13Ralph2 when administered to an animal, outside of the subject; and
CC administering the vaccine to the subject to stimulate an immune response
CC against IL-13Ralph2 in the subject. Also described: (1) a composition
CC for stimulating the immune response against IL-13Ralph2 when
CC administered to an animal, comprising an isolated gene that can stimulate
CC an immune response against IL-13Ralph2 when administered to the animal,
CC and carrier; (2) a method for directing an antibody to cells expressing
CC IL-13Ralph2 in a subject by formulating a pharmaceutical composition
CC outside of a subject, where the composition comprises an antibody that
CC specifically binds IL-13Ralph2 and a carrier; and administering the
CC pharmaceutical composition to the subject to allow the antibody to
CC specifically bind to the cells expressing IL-13Ralph2 in the subject;
CC and (3) a pharmaceutical composition comprising the antibody that
CC specifically binds IL-13Ralph2 and a carrier. The methods and
CC composition are useful for stimulating an immune response against IL-
CC 13Ralph2 in a subject having or at risk for developing a disease having
CC cells expressing IL-13Ralph2. The present sequence represents an
CC immunogenic peptide for IL-13Ralph2, which is used in an example from
CC the present invention.
CC
XX Sequence 10 AA;

Query Match 30.3%; Score 30; DB 7; Length 10;
Best Local Similarity 57.1%; Pred. No. 7.5e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSDGLMN 7
| | | | | : |
| | | | | : |
DB 2 SDDGLMS 8

RESULT 33
AAE10548
ID AAE10548 standard; peptide; 12 AA.
XX
AC AAE10548;
XX
DT 10-DEC-2001 (first entry)
XX
DE llama species antibody VHH CDR3 #17.

XX
KM llama antibody; camelid; anorectic; heavy chain variable domain; VHH;
KM human dietary enzyme inhibitor; medicament; human pancreatic lipase; HPL;
KM food; human gastric lipase; HGL; cosmetic control; body weight;
KM complementarity determining region 3; CDR3.
XX
OS Lama sp.
PN EP1134231-A1.
XX
PD 19-SEP-2001.
XX
PF 20-FEB-2001; 2001EP-00200703.
XX
PR 14-MAR-2000; 2000EP-00200930.
XX
PA (UNITL) UNILEVER NV.
PA (UNITL) UNILEVER PLC.
PI Bezemer S, Van De Burg M, De Haard UW, Tareilus E;
XX
DR WPI; 2001-572718/65.

XX
PT New antibody or its fragments for inhibiting human dietary enzymes,
PT useful for cosmetic control of body weight of human beings, comprises
PT heavy chain variable domain derived from immunoglobulin naturally devoid
PT of light chains.
XX
XX
XX Claim 4; Page 29; 37p; English.

XX
PS The patent discloses antibodies or their fragments comprising a heavy
CC chain variable domain (VHH) derived from an immunoglobulin naturally
CC devoid of light chains specific for inhibiting human dietary enzymes. The
CC antibodies of the invention are useful for the preparation of medicaments
CC or food for inhibiting the activity of one or more human dietary enzymes
CC especially human pancreatic lipase (HPL) or human gastric lipase (HGL)
CC which are useful for the cosmetic control of body weight of human beings.
CC The present peptide sequence is a complementarity determining region 3
CC (CDR3) of llama species (camelid) antibody VHH region
XX
XX Sequence 12 AA;

Query Match 30.3%; Score 30; DB 4; Length 12;
Best Local Similarity 62.5%; Pred. No. 9.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSDGLMNN 8
| | | | | : |
| | | | | : |
DB 4 SYDGSWMD 11

RESULT 34
ABB80837
ID ABB80837 standard; protein; 12 AA.

```

XX AC ABB80837;
XX
XX DT 23-SEP-2002 (first entry)
XX
XX DE Heparin binding laminin alpha4 protein fragment.
XX
XX KW Laminin alpha4; heparin; signal transduction; transgenic; cancer;
XX KW angiogenesis; cytosstatic; vasotropic; gene therapy; mouse.
XX
XX OS Mus musculus.
XX
XX PN WO200248349-A1.
XX
XX PD 20-JUN-2002.
XX
XX PF 10-JUL-2001; 2001WO-JP005976.
XX
XX PR 12-DEC-2000; 2000JP-00376899.
XX
XX PA (KYOW ) KYOWA HAKKO KOGYO KK.
XX
XX PI Kitagawa Y, Shitara K, Ohki Y;
XX
XX DR WPI; 2002-490604/52.
XX
XX PT Polypeptides binding strongly to heparin for treatment of cancer and
XX PT angiogenesis-related disorders.
XX
XX PS Example; Page 106; 129pp; Japanese.
XX
XX CC The invention relates to a polypeptide forming part of the laminin alpha4
XX CC chain G domain, having molecular weight below 108 kDa, and capable of
XX CC binding to heparin. The polypeptide can be used to inhibit heparin-
XX CC binding signal transduction from cells. The polypeptide, encoding DNA,
XX CC non-human transgenic animals and transgenic plants containing the
XX CC recombinant DNA can all be used for the treatment and prevention of
XX CC cancer and angiogenesis-related diseases. The present sequence represents
XX CC a fragment of the heparin binding mouse laminin alpha4 protein
XX
XX SQ Sequence 12 AA;
XX
XX Query Match 30.3%; Score 30; DB 5; Length 12;
XX Best Local Similarity 57.1%; Pred. No. 9.1e+02;
XX Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2 SDGLMNN 8
XX Db 6 NDGLMHD 12
XX
XX RESULT 35
XX AAB73315
XX ID AAB73315 standard; peptide; 14 AA.
XX
XX AC AAB73315;
XX
XX DT 22-MAY-2001 (first entry)
XX
XX DE Human PC-LECTIN extracellular domain peptide, SEQ ID NO:25.
XX
XX KW Human; PC-LECTIN; C-type lectin; transmembrane antigen; normal testis;
XX KW layilin homologue; prostate cancer antigen; overexpression;
XX KW androgen-dependent prostate cancer; diagnosis; prognosis;
XX KW extracellular domain; antigenic; antibody.
XX
XX OS Homo sapiens.
XX
XX PN WO200112811-A1.
XX
XX PD 22-FEB-2001.
XX
XX PF 11-AUG-2000; 2000WO-US022065.

```

```

XX PR 12-AUG-1999; 99US-0148935P.
XX
XX XX (USOG-) UROGENESYS INC.
XX
XX PA
XX
XX PF Afar DEH, Hubert RS, Jakobovits A, Raitano AB;
XX
XX DR WPI; 2001-211222/21.
XX
XX PT New PC-LECTIN polynucleotide encoding a transmembrane antigen over
XX PT expressed in human prostate cancer, useful for the prognosis, diagnosis
XX PT and treatment of prostate cancer.
XX
XX PS Claim 4; Page 79; 116pp; English.
XX
XX CC The invention relates to a novel human C-type lectin transmembrane
XX CC antigen, PC-LECTIN (AAB73309) and cDNA encoding it (AAB76004). The
XX CC expression of the human PC-LECTIN gene is normally restricted to the
XX CC testis, but is highly overexpressed in prostate cancer. PC-LECTIN
XX CC expression is higher in androgen-dependent prostate tumours compared with
XX CC androgen-independent prostate tumours, and expression is therefore likely
XX CC to be dependent on the presence of androgen. Human PC-LECTIN therefore
XX CC represents a diagnostic and therapeutic target for prostate cancer.
XX CC particularly androgen-dependent prostate cancer. Human PC-LECTIN exhibits
XX CC homology to hamster layilin (44.9% identity over a 265 residue overlap),
XX CC but is not thought to be the human orthologue of layilin, as diverges
XX CC significantly in a key functional domain proposed for the layilin
XX CC protein. Human PC-LECTIN or an immunogenic portion thereof, a vector
XX CC encoding PC-LECTIN, a PC-LECTIN antisense nucleotide, a PC-LECTIN
XX CC nucleotide-targeted ribozyme, or an anti-PC-LECTIN antibody may be used
XX CC to prepare a composition for treating a patient with a cancer,
XX CC particularly prostate cancer, but also breast, bladder, lung, bone,
XX CC colon, pancreatic, testicular, cervical or ovarian cancers that express
XX CC PC-LECTIN. PC-LECTIN proteins are also useful for diagnosing the presence
XX CC of cancer. PC-LECTIN antibodies and nucleotides are useful in the
XX CC treatment (e.g., antisense therapy), diagnosis and/or prognosis of
XX CC prostate cancer and other PC-LECTIN-expressing cancers. PC-LECTIN
XX CC antibodies may also be used as drug targeting agents. The PC-LECTIN
XX CC nucleotides and proteins may additionally be used in drug discovery to
XX CC identify molecules that modulate PC-LECTIN function or expression. The
XX CC present sequence represents a human PC-LECTIN extracellular domain
XX CC peptide used to raise anti-PC-lectin antibodies
XX
XX SQ Sequence 14 AA;
XX
XX Query Match 30.3%; Score 30; DB 4; Length 14;
XX Best Local Similarity 55.6%; Pred. No. 1.1e+03;
XX Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX QY 4 GLMNNNQTO 12
XX Db 1 GLMNNNQTO 9
XX
XX RESULT 36
XX ABJ00795
XX ID ABJ00795 standard; peptide; 14 AA.
XX
XX AC ABJ00795;
XX
XX DT 05-SEP-2002 (first entry)
XX
XX DE B lymphocyte affinity maturation library peptide #69.
XX
XX KW B lymphocyte stimulator protein binding protein; BVS; immune disease;
XX KW allergy; proliferative disease; infectious disease; arteriosclerosis;
XX KW inflammatory disorder; hypergammaglobulinemia; blood clotting;
XX KW ischaemia; graft-versus-host disease; neurodegenerative disease;
XX KW immunosuppressive; nephrotropic; antineumatic; antiarthritic;
XX KW neuroprotective; cytosstatic; immunostimulant; antitumour; anti-HIV;
XX KW antiaesthetic; antiallergic; thyromimetic; antianemic; haemostatic;
XX KW dermatological; antinflammatory; cardiant; ophthalmological; uropathic;
XX KW antidiabetic; antithyroid; antidepressant; hepatotropic.

```

XX OS Unidentified.
 XX PN WO200216411-A2.
 XX PD 28-FEB-2002.
 XX PF 17-AUG-2001; 2001WO-US025850.
 XX PR 18-AUG-2000; 2000US-0226700P.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PI Beltzer JP, Potter DM, Fleming TL, Rosen CA;
 XX DR WPI; 2002-499775/53.
 XX PT The treatment of various diseases e.g. rheumatoid arthritis, comprises
 PT administering B lymphocyte stimulator binding polypeptide.
 XX PS Claim 70; Page 215; 387pp; English.
 XX CC The present invention relates to the treatment, prevention or
 CC amelioration of a disease or disorder associated with: aberrant B
 CC lymphocyte stimulator (Blys), Blys receptor expression or activity; cells
 CC of haematopoietic origin; or proliferative disease; and reducing,
 CC inhibiting or stimulating immunoglobulin production, B cell proliferation
 CC and graft rejection involving administration of Blys binding polypeptide.
 CC The Blys binding polypeptides are used in the treatment, prevention or
 CC amelioration of diseases such as immune system diseases, proliferative or
 CC diseases, diseases of cells of haematopoietic origin, graft rejection,
 CC allergies, infectious diseases, arteriosclerosis, inflammatory disorders,
 CC hyperimmunoglobulinaemia, blood clotting disorders, ischaemia, and
 CC neurodegenerative diseases. The present sequence is a B lymphocyte
 CC stimulator protein binding peptide
 XX SQ Sequence 14 AA;
 OY Query Match 30.3%; Score 30; DB 5; Length 14;
 DB Best Local Similarity 38.5%; Pred. No. 1.1e+03;
 Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 5 LMNNQOTQLFLEH 17
 2 IWDPLTKMLPH 14
 RESULT 37
 ABG33656
 ID ABG33656 standard; peptide; 14 AA.
 XX AC ABG33656;
 XX DT 15-JUL-2002 (first entry)
 XX DE B Lymphocyte Stimulator (Blys) binding peptide #232.
 XX KM B Lymphocyte Stimulator protein; B Lymphocyte Stimulator binding peptide;
 KM Blys; biological fluid; serum; plasma; lymph; blood; urine; spinal fluid;
 KM synovial fluid; saliva; mucus.
 XX OS Synthetic.
 XX PN WO200216412-A2.
 XX PD 28-FEB-2002.
 XX PF 17-AUG-2001; 2001WO-US025891.
 XX PR 18-AUG-2000; 2000US-0226489P.
 XX PA (DYAX-) DYAX CORP.

PI Beltzer JP, Potter MD, Fleming TJ, Ladner RC;
 XX DR WPI; 2002-351647/38.
 XX PT New B-lymphocyte stimulator binding polypeptide useful in detecting or
 PT isolating Blys or Blys-like polypeptide comprises a specified amino acid
 XX sequence.
 XX PS Claim 6; Page 89; 269pp; English.
 XX CC The invention relates to a B lymphocyte stimulator (Blys) binding
 CC polypeptide. Blys binding peptides bind Blys or Blys-like proteins
 CC reversibly or irreversibly. The binding peptides are used in detection,
 CC isolation and/or purification of Blys in a solution such as water or a
 CC buffer solution, as well as any fluid and/or cell obtained from an
 CC individual biological fluid, body tissue, body cell, cell line, tissue
 CC culture or other source containing Blys or Blys-like polypeptides. The
 CC biological fluids include sera, plasma, lymph, blood, blood fraction,
 CC urine, synovial fluid, spinal fluid, saliva and mucous. Sequences
 CC ABG33406-33415, ABG33423-33575, ABG33588-33846, ABG33848-33850 and
 CC ABG33852-33862 represent Blys binding peptides of the invention
 XX SQ Sequence 14 AA;
 OY Query Match 30.3%; Score 30; DB 5; Length 14;
 DB Best Local Similarity 38.5%; Pred. No. 1.1e+03;
 Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 5 LMNNQOTQLFLEH 17
 2 IWDPLTKMLPH 14
 RESULT 38
 AAR85392
 ID AAR85392 standard; peptide; 15 AA.
 XX AC AAR85392;
 XX DT 13-JUN-1996 (first entry)
 XX DE HTLV-IIIB envelope glycoprotein antigenic peptide, SP-15.
 XX KM Human immunodeficiency virus; HIV; HTLV-IIIB; envelope; glycoprotein;
 KM hydrophilic; immunisation; antibody production; fusion peptide; SP-1;
 KM conjugate; carrier.
 XX OS Synthetic.
 XX PN WO9529700-A1.
 XX PD 09-NOV-1995.
 XX PF 28-APR-1995; 95WO-US005465.
 XX PR 29-APR-1994; 94US-00235305.
 XX PA (UYDU-) UNIV DUKE.
 XX PI Haynes BF, Parker TJ;
 XX DR WPI; 1995-392926/50.
 XX PT New peptide(s) corresponding to HIV sequences - used for inducing
 PT protective immunity to HIV and in the treatment of e.g. auto-immune
 PT disease, infectious disease or tumours.
 XX PS Example 1; Page 42; 128pp; English.
 CC AAR85388-R85393 synthetic peptides with hydrophilic amino acid sequences
 CC of the HIV isolate HTLV-IIIB envelope protein. These peptides can be
 CC covalently linked to a carrier molecule. The conjugate produced is
 CC capable of inducing the production of high titres of protective, type-

CC specific, anti-HIV antibodies. Helper T cells and cytotoxic T cells are
 CC also activated by the peptide immunogens
 XX
 SQ Sequence 15 AA;

Query Match 30.3%; Score 30; DB 2; Length 15;
 Best Local Similarity 41.7%; Pred. No. 1.2e+03;
 Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
 QY 3 DGLMNNQTOLE 14
 ||:||||:
 Db 4 DGSNSNSESEIF 15

RESULT 39
 AAW19031
 ID AAW19031 standard; peptide; 15 AA.

AC AAW19031;
 XX
 DT 16-JAN-1998 (first entry)
 XX
 DE SP-15 region of HIV envelope glycoprotein.

XX Hydrophilic; antigenic determinant; HIV; envelope; glycoprotein; env; gp;
 KW recognition; B lymphocyte; type specific; antibody; vaccine; protection;
 KM immune response; infection; neutralisation; SP-15 region.

XX Human immunodeficiency virus.

XX MO9714436-A1.

XX 24-APR-1997.

XX 18-OCT-1996; 96MO-US016911.

XX 20-OCT-1995; 95US-00546515.

XX 09-FEB-1996; 96US-00599266.

XX (UYDU-) UNIV DUKE.

XX Haynes BF, Palker TJ;

XX WPI; 1997-244862/22.

XX Synthetic human immunodeficiency virus vaccine - comprising hydrophilic
 PT peptide corresponding to at least 1 antigenic determinant of envelope
 PT glyco:protein recognised by B lymphocytes.

XX Example 1; Page 37; 104pp; English.

XX An essentially pure hydrophilic peptide, comprising at least 1 antigenic
 CC determinant of human immunodeficiency virus (HIV) envelope (env)
 CC glycoprotein (gp) recognised by B lymphocytes, i.e. the present sequence,
 CC when covalently linked to a carrier molecule, induces the production of
 CC high titres of protective, type specific anti-HIV antibodies (Ab) in a
 CC mammal. The peptide can be used in vaccines for producing a protective
 CC immune response to HIV infection, while a HIV neutralising Ab can be
 CC induced in a primate by administering a composition comprising HIV env
 CC peptides that disrupt gp120/gp41 interactions

XX Sequence 15 AA;

Query Match 30.3%; Score 30; DB 2; Length 15;
 Best Local Similarity 41.7%; Pred. No. 1.2e+03;
 Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 3 DGLMNNQTOLE 14
 ||:||||:
 Db 4 DGSNSNSESEIF 15

RESULT 40

ADL25934
 ID ADL25934 standard; peptide; 15 AA.

XX ADL25934;

XX 17-JUN-2004 (first entry)

XX Synthetic peptide A121 derived from a conserved region of HCV.

XX HCV; hepatitis C virus; virucide; vaccine; MHC; HLA;

XX major histocompatibility complex; human leukocyte antigen.

XX Synthetic.

XX WO2004024182-A2.

XX 25-MAR-2004.

XX 27-AUG-2003; 2003WO-EP009482.

XX 13-SEP-2002; 2002AT-00001376.

XX 27-FEB-2003; 2003WO-EP002005.

XX 11-JUL-2003; 2003EP-00450171.

XX (INTE-) INTERCELL AG.

XX Buschle M, Habel A, Klade C, Mattner F, Otava O, Vytvytska O;

XX Zauner W, Zinke S, Kiriappos H;

XX WPI; 2004-269899/25.

XX Isolating Hepatitis C Virus peptides (HPS) which have a binding capacity
 PT to a MHC/HLA molecule or a complex comprising the HCV-peptide and the
 PT molecule by separating the complex from the HCV-peptides which do not
 PT bind to the molecule.

XX Example 1; Page 30; 73pp; English.

XX The invention relates to a novel method for isolating Hepatitis C Virus
 CC (HCV) peptides (HPS). The method of the invention has virucide activity,
 CC and may be useful in producing a vaccine. The method is useful for
 CC isolating Hepatitis C Virus peptides (HPS) which have a binding capacity
 CC to a MHC/HLA molecule or a complex comprising the HCV-peptide and the
 CC MHC/HLA molecule for preparing a vaccine against HCV infection. The T
 CC cells, a T cell clone or a T cell population or preparation is useful for
 CC identifying heteroclitic epitopes or for preparing a composition for
 CC treating HCV infection. The present sequence represents a synthetic
 CC peptide derived from a conserved region of HCV.

XX Sequence 15 AA;

Query Match 30.3%; Score 30; DB 8; Length 15;
 Best Local Similarity 41.7%; Pred. No. 1.2e+03;
 Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 SDGLMNNQTOLE 13
 ::||:|:
 Db 1 TNGSMHINRTAL 12

Search completed: December 30, 2004, 16:28:14
 Job time : 160 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 16:04:21 ; Search time 38 Seconds

(without alignments)
45.576 Million cell updates/sec

Title: US-10-718-321-1

Perfect score: 99
Sequence: 1 SSDGLMNNNOTQLFLHS 18Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 150 summaries

Database :

1: PIR 79:*\n2: PIR1:*\n3: PIR2:*\n4: PIR3:*\n4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84.8	451	2	S71754	cellular hepatitis
2	49.5	267	2	AB2541	hypothetical prote
3	49.5	338	2	S75217	N-acetylmuramoyl-L
4	48.5	416	2	S09334	transcobalamin I p
5	48.5	1254	1	J01978	structural polypor
6	48.5	1254	1	VHMVVE	structural polypor
7	48.5	1254	1	VHMVVE	structural polypor
8	48.5	1254	1	VHMVVE	structural polypor
9	48.5	1255	1	B44213	structural polypor
10	47.5	1404	2	T06663	DNA (cytosine-5'-)
11	46.5	508	1	S46008	probable carboxyph
12	45.5	230	2	AB6445	unknown protein [i
13	45.5	301	1	LNRT2	hepatic lectin 2 -
14	45.5	1464	1	S29159	glutamate receptor
15	45.5	1464	1	A43274	N-methyl-D-asparta
16	44.4	589	2	A29476	muscarinic acetylch
17	44.4	589	2	B29514	muscarinic acetylch
18	44.4	590	2	S10138	muscarinic acetylch
19	44.4	590	2	S10138	muscarinic acetylch
20	44.4	590	2	S10138	muscarinic acetylch
21	44.4	639	2	A55019	muscarinic acetylch
22	44.4	973	2	T35238	probable secreted
23	44.4	1042	1	GBRCE	beta-galactosidase
24	44.4	1042	2	B85968	evolved beta-D-gal
25	44.4	1042	2	P81133	evolved beta-D-gal
26	44.4	1084	2	C82931	hypothetical prote
27	43.4	254	2	S76938	hypothetical prote
28	42.4	153	2	AG2598	conserved hypothet
29	42.4	153	2	H97380	phng protein [limp

30	42	42.4	269	2	H65067	hypothetical prote
31	42	42.4	269	2	H91091	probable sensory t
32	42	42.4	269	2	C85937	probable sensory t
33	42	42.4	278	1	S73153	hypothetical prote
34	42	42.4	326	1	VGXR37	glycoprotein VP7 p
35	42	42.4	400	2	A97535	ispd/ispf bifuncti
36	42	42.4	400	2	AC2754	ispd/ispf bifuncti
37	42	42.4	457	2	T05439	hypothetical prote
38	42	42.4	620	2	A58932	cytochrome C-type
39	42	42.4	1014	2	S75724	hypothetical prote
40	42	42.4	1161	2	S72645	pili protein - ps
41	42	42.4	1161	2	D83076	type 4 fibrillar bi
42	41.5	41.9	786	2	T39034	probable tRNA ligase
43	41.5	41.9	1254	1	J01979	structural polypor
44	41.5	41.4	266	2	T07363	hypothetical prote
45	41	41.4	326	1	VGXR38	glycoprotein VP7 p
46	41	41.4	326	1	VGXR38	glycoprotein VP7 p
47	41	41.4	326	1	VGXR38	glycoprotein VP7 p
48	41	41.4	326	1	VGXR38	glycoprotein VP7 p
49	41	41.4	326	1	VGXR38	glycoprotein VP7 p
50	41	41.4	486	2	AC2474	hypothetical prote
51	41	41.4	486	2	AC2474	hypothetical prote
52	41	41.4	486	2	AC2474	hypothetical prote
53	41	41.4	569	2	C86334	probable membrane
54	41	41.4	656	2	E75468	hypothetical prote
55	41	41.4	955	2	A60990	alpha-amylose (EC
56	41	41.4	1455	1	A48925	mannose receptor p
57	41	41.4	1464	2	S47555	N-methyl-D-asparta
58	40.5	40.9	226	2	T27843	hypothetical prote
59	40.5	40.9	720	2	B15540	glycogen operon pr
60	40	40.4	260	1	AC2297	hypothetical prote
61	40	40.4	260	1	AC2297	protein phosphatas
62	40	40.4	300	2	T24292	hypothetical prote
63	40	40.4	312	2	S51440	hypothetical prote
64	40	40.4	390	2	AT3359	2-C-methyl-D-eryth
65	40	40.4	483	2	D71439	probable Beta-Amyl
66	40	40.4	483	2	T21327	hypothetical prote
67	40	40.4	489	1	S16390	mitochondrial proc
68	40	40.4	498	2	H85190	probable beta-amy1
69	40	40.4	532	1	CPBYV	carboxypeptidase C
70	40	40.4	548	2	T52556	beta-amylose (EC 3
71	40	40.4	576	2	T38666	probable trp-asp r
72	40	40.4	583	2	T00091	xgpc protein - St
73	40	40.4	645	2	S52047	desoxyribodipyrimid
74	40	40.4	645	2	T32820	hypothetical prote
75	40	40.4	707	2	T22127	hypothetical prote
76	40	40.4	891	2	T30812	ubiquitin-protein
77	40	40.4	1166	2	T29009	hypothetical prote
78	40	40.4	1171	2	T00380	hypothetical prote
79	40	40.4	1239	1	VHMVEE	KIAA0637 protein -
80	40	40.4	1240	1	VHMVEE	structural polypor
81	40	40.4	1241	2	S26373	genome polypor
82	40	40.4	1242	2	A56605	structural polypor
83	40	40.4	1242	2	S72350	structural polypor
84	40	40.4	1665	2	T29008	hypothetical prote
85	40	40.4	2004	2	AC0314	hypothetical prote
86	40	40.4	2149	2	C96695	probable membrane
87	39.5	39.9	405	2	T41712	ribulose biphosph
88	39.5	39.9	732	2	T00130	hypothetical prote
89	39	39.4	130	2	S11562	trehalose synthase
90	39	39.4	232	2	AH0252	hypothetical prote
91	39	39.4	268	2	C90484	conserved hypothet
92	39	39.4	291	2	G82136	hypothetical prote
93	39	39.4	297	2	D64206	hypothetical prote
94	39	39.4	326	1	VGXR4	glycoprotein VP7 p
95	39	39.4	326	1	VGXR4	glycoprotein VP7 p
96	39	39.4	402	2	T29478	glycoprotein VP7 p
97	39	39.4	407	2	G71414	glycoprotein VP7 p
98	39	39.4	439	1	H82875	hydroxymandelonitr
99	39	39.4	455	1	T21089	oligopeptide trans
100	39	39.4	466	2	T21089	acid phosphatase (
101	39	39.4	476	2	H82085	ORF MSV235 probabl
102	39	39.4	482	2	T49079	sodium/alanine sym
						serine-type carbox

103	39	39.4	484	2	A40774	phosphocholine-bin
104	39	39.4	487	2	T49080	serine-type carbox
105	39	39.4	489	2	S50396	hypothetical prote
106	39	39.4	511	2	P66752	protein phosphatas
107	39	39.4	516	2	A56590	ecdysteroid recept
108	39	39.4	543	2	U04812	hyaluronan synthas
109	39	39.4	549	2	S44168	alpha-glucosidase
110	39	39.4	588	2	A43740	Dg42 protein - Afr
111	39	39.4	725	2	T03219	G-quartet DNA bind
112	39	39.4	736	2	B28341	photosystem I P700
113	39	39.4	747	2	T33746	hypothetical prote
114	39	39.4	810	2	B16639	vib4 protein prec
115	39	39.4	822	2	T33163	hypothetical prote
116	39	39.4	1137	2	S13759	morphogenesis-rela
117	39	39.4	1481	2	S28669	pullulanase (EC 3.
118	39	39.4	1707	2	AH2085	two-component hybr
119	38.5	38.9	646	2	S38819	plasma protein S -
120	38.5	38.9	1374	2	AE3259	extracellular seri
121	38	38.4	88	2	S77431	hypothetical prote
122	38	38.4	182	2	T27367	hypothetical prote
123	38	38.4	190	2	A33619	GTP-binding protei
124	38	38.4	229	2	A05198	hypothetical prote
125	38	38.4	230	2	B82252	RafE-related prote
126	38	38.4	231	2	AE2536	hypothetical prote
127	38	38.4	232	2	B82935	deoxyguanosine kin
128	38	38.4	246	2	AE2366	ABC transporter AT
129	38	38.4	252	2	PC1140	cellulase (EC 3.2.
130	38	38.4	263	2	A11815	indazoleglycerol-
131	38	38.4	264	2	T49055	serine carboxypept
132	38	38.4	270	2	B84578	probable benesenc
133	38	38.4	295	2	B82502	transcription regu
134	38	38.4	301	2	S13165	asialoglycoprotein
135	38	38.4	307	2	D97701	hypothetical prote
136	38	38.4	316	2	AG0640	flagellar protein
137	38	38.4	316	2	JC3930	flagellar protein
138	38	38.4	320	2	JC1255	T1811 protein - ra
139	38	38.4	326	1	VGXRT6	glycoprotein VP7 p
140	38	38.4	326	1	VGXRT6	glycoprotein VP7 p
141	38	38.4	326	1	VGXRT6	glycoprotein VP7 p
142	38	38.4	326	1	VGXRT6	glycoprotein VP7 p
143	38	38.4	326	1	VGXRT6	glycoprotein VP7 p
144	38	38.4	334	2	D65097	YJ1R protein - Bsc
145	38	38.4	334	2	A98125	hypothetical prote
146	38	38.4	334	2	B85969	hypothetical prote
147	38	38.4	340	2	S75099	hypothetical prote
148	38	38.4	340	2	T16733	hypothetical prote
149	38	38.4	350	2	T00892	hypothetical prote
150	38	38.4	357	2	E17708	hypothetical prote

ALIGNMENTS

RESULT 1
S71754
cellular hepatitis A receptor HAVcr-1 precursor - green monkey
N:Alternate names: surface glycoprotein
C:Species: Cercopithecus aethiops (green monkey, grivet)
C>Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 09-Jul-2004
C:Accession: S71754
R:Kaplan, G.; Totenka, A.; Thompson, P.; Akatsuka, T.; Moritsugu, Y.; Feinstein, S.M.
EMBO J. 15, 4282-4296, 1996
A:Title: Identification of a surface glycoprotein on African green monkey kidney cells
A:Reference number: S71754; MUID:97015129; PMID:8861957
A:Accession: S71754
A:Molecule type: mRNA
A:Residues: 1-451 <KAP>
A:Cross-references: UNIPROT:Q95144; EMBL:X98252; NID:G1526573; PID:e247449; PID:G1526574
A:Experimental source: kidney
C:Keywords: glycoprotein
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-451/Product: cellular hepatitis A receptor HAVcr-1 #status predicted <MAT>

Query Match 84.8%; Score 84; DB 2; Length 451;
Best Local Similarity 88.9%; Pred. No. 1,2e-05;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db 1 SSDGMMNNOTQLEHS 18
344 SSDGMMNNOTQLEHS 361

RESULT 2
AB2541
hypothetical protein alr7613 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120bet
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AB2541
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuchih,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anal
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AB2541
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-267 <KUR>
A:Cross-references: UNIPROT:Q82598; GB:AB003602; PIDN:BA877256.1; PID:G17134698; GSPDB:GR
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr7613
A:Genome: plasmid

Query Match 49.5%; Score 49; DB 2; Length 267;
Best Local Similarity 57.1%; Pred. No. 2.6;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 2 SDGMMNNOTQLE 15
190 SEGMMNNOTQDFV 203

RESULT 3
S75217
N-acetyl[muramoyl]-L-alanine amidase (EC 3.5.1.28) 1 - Synechocystis sp. (strain PCC 6803)
N:Alternate names: protein slr1910
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S75217
R:Kaneko, T.; Sato, S.; Kocani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
A:Reference number: S74322; MUID:97061201; PMID:8995231
A:Accession: S75217
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-338 <KAN>
A:Cross-references: UNIPROT:P71105; EMBL:D90903; GB:AB001339; NID:G1652127; PIDN:BAA17131
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Gene: amIA-1
C:Keywords: hydrolase

Query Match 49.5%; Score 49; DB 2; Length 338;
Best Local Similarity 61.5%; Pred. No. 3.4;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Db 4 GLMMNNOTQLEFL 16
95 GRMMNNSTRIVLE 107

```
RESULT 4
509334
transcobalamin I precursor - pig (fragment)
N/Alternate names: haptocorrin
C/Species: Sus scrofa domestica (domestic pig)
C/Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jul-2004
C/Accession: S09334
R/Helit, J.E.; Seetharam, B.; Levkam, J.; Alpers, D.H.
Eur J. Biochem. 189, 125-130, 1990
A/Title: Isolation and characterization of a cDNA encoding porcine gastric haptocorrin.
A/Reference number: S09334; MUID:90235848; PMID:2331993
A/Accession: S09334
A/Molecule type: mRNA
A/Residues: 1-416 <HEM>
A/Cross-references: UNIPROT:P17630; EMBL:X52566; NID:g1963; PIDN:CAA36800.1; PID:g1964
C/Superfamily: gastric intrinsic factor
C/Keywords: glycoprotein
F/1-24/Domain: signal sequence (fragment) #status predicted <SIG>
F/25-416/Product: haptocorrin #status predicted <MAT>

Query Match          48.5%; Score 48; DB 2; Length 416;
Best Local Similarity 44.4%; Pred. No. 6.3;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Cy 1 SSDGLMNNNOTOLFLEHS 18
Db 378 SVGGIMANNERTYWEHS 395
```

```
RESULT 5
J01978
structural polyprotein - Venezuelan equine encephalitis virus (subtype II, strain Evergl)
N/Contains: 6K protein; coat protein; membrane glycoprotein E1; membrane glycoprotein E2
C/Species: Venezuelan equine encephalitis virus
C/Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C/Accession: J01978
R/Sheider, J.M.; Kinney, R.M.; Tsuchiya, K.R.; Trent, D.W.
J. Gen. Virol. 74, 519-523, 1993
A/Title: Molecular evidence that epizootic Venezuelan equine encephalitis (VEE) T-AB vir
A/Reference number: J01978; MUID:93187617; PMID:8445371
A/Accession: J01978
A/Molecule type: mRNA
A/Residues: 1-1254 <SNE>
A/Cross-references: UNIPROT:P36330; GB:L04598; NID:s290612; PIDN:AAA42984.1; PID:s290614
C/Superfamily: togavirus structural polyprotein
C/Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F/1-274/Product: coat protein #status predicted <CTP>
F/275-333/Product: membrane glycoprotein E3 #status predicted <MG3>
F/334-756/Product: membrane glycoprotein E2 #status predicted <MG2>
F/701-721/Domain: transmembrane #status predicted <TM1>
F/757-812/Product: 6K protein #status predicted <KPe>
F/794-813/Domain: transmembrane #status predicted <TM2>
F/813-1254/Product: membrane glycoprotein E1 #status predicted <MG1>
F/1231-1248/Domain: transmembrane #status predicted <TM3>
F/47,285,545,651,946/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match          48.5%; Score 48; DB 1; Length 1254;
Best Local Similarity 57.9%; Pred. No. 21;
Matches 11; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

Cy 1 SSDGLMNNNO---TOLFL 15
Db 762 SLDLHMNNNQMFMTOLLI 780
```

```
RESULT 6
VHMVVE
structural polyprotein - Venezuelan equine encephalitis virus (strain TC-83)
N/Contains: 6K peptide; coat protein; membrane glycoprotein E1; membrane glycoprotein E2
C/Species: Venezuelan equine encephalitis virus
C/Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 09-Jul-2004
C/Accession: A27871
R/Johnson, B.J.B.; Kinney, R.M.; Kost, C.L.; Trent, D.W.
J. Gen. Virol. 67, 1951-1960, 1986
A/Title: Molecular determinants of alphavirus neurovirulence: nucleotide and deduced pr
A/Reference number: A27871; MUID:86306669; PMID:3755750
A/Accession: A27871
A/Molecule type: mRNA
A/Residues: 1-1254 <JOH>
A/Cross-references: UNIPROT:P05674; GB:X04368; NID:958621; PIDN:CAA27883.1; PID:958622
A/Note: the authors translated the codon UUC for residues 19 and 27 as Ile and Acc for r
C/Superfamily: togavirus structural polyprotein
C/Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F/1-75/Domain: transmembrane #status predicted <TM1>
F/276-334/Product: coat protein #status predicted <CTP>
F/335-757/Product: membrane glycoprotein E3 #status predicted <GP3>
F/758-812/Domain: transmembrane #status predicted <TM1>
F/813-1254/Product: membrane glycoprotein E2 #status predicted <GP2>
F/1231-1247/Domain: transmembrane #status predicted <TM4>
F/47,286,546,625,652,946/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match          48.5%; Score 48; DB 1; Length 1254;
Best Local Similarity 50.0%; Pred. No. 21;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Cy 1 SSDGLMNNNOTOLFLE 16
Db 763 SLDLHMNNNQMFMTIQ 778
```

```
RESULT 7
VHMVVT
structural polyprotein - Venezuelan equine encephalitis virus (strain TRD)
N/Contains: 6K protein; coat protein; membrane glycoprotein E1; membrane glycoprotein E2
C/Species: Venezuelan equine encephalitis virus
C/Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C/Accession: B31467; A47612
R/Kinney, R.M.; Johnson, B.J.B.; Welch, J.B.; Tsuchiya, K.R.; Trent, D.W.
Virology 170, 19-30, 1989
A/Title: The full-length nucleotide sequences of the virulent Trinidad donkey strain of
A/Reference number: A31467; MUID:89243175; PMID:2524126
A/Accession: B31467
A/Molecule type: mRNA
A/Residues: 1-1254 <RI>
A/Cross-references: UNIPROT:P09592; GB:J04332; NID:g323708; PIDN:AA802519.1; PID:g323710
R/Kinney, R.M.; Johnson, B.J.B.; Brown, V.L.; Trent, D.W.
Virology 152, 400-413, 1986
A/Title: Nucleotide sequence of the 26 S mRNA of the virulent Trinidad donkey strain of
A/Reference number: A47612; MUID:86263392; PMID:3088830
A/Accession: A47612
A/Molecule type: mRNA
A/Residues: 1-542, 'K', 544-810, 'P', 812-1254 <RI2>
A/Cross-references: GB:L01442
C/Superfamily: togavirus structural polyprotein
C/Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F/1-275/Product: coat protein #status predicted <CTP>
F/276-334/Product: membrane glycoprotein E3 #status predicted <MG3>
F/335-757/Product: membrane glycoprotein E2 #status predicted <MG2>
F/701-718/Domain: transmembrane #status predicted <TM1>
F/758-812/Product: 6K protein #status predicted <KPe>
F/774-790/Domain: transmembrane #status predicted <TM2>
F/795-813/Domain: transmembrane #status predicted <TM3>
F/813-1254/Product: membrane glycoprotein E1 #status predicted <MG1>
F/1231-1248/Domain: transmembrane #status predicted <TM4>
F/47,286,546,652,946/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match          48.5%; Score 48; DB 1; Length 1254;
Best Local Similarity 50.0%; Pred. No. 21;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Cy 1 SSDGLMNNNOTOLFLE 16
Db 763 SLDLHMNNNQMFMTIQ 778
```

RESULT 8

B44213

structural polypeptide - Venezuelan equine encephalitis virus (strain P676)
N:Contains: 6K protein; coat protein; membrane glycoprotein E1; membrane glycoprotein E2
C:Species: Venezuelan equine encephalitis virus
C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C/Accession: B44213

R:Kinney, R.M.; Teuchiya, K.R.; Snider, J.M.; Trent, D.W.
Virology 191, 569-580, 1992

A>Title: Genetic evidence that epizootic Venezuelan equine encephalitis (VEE) viruses ma
A/Reference number: A44213; MUID:93079859; PMID:1448915
A/Accession type: genomic RNA

A/Molecule type: genomic RNA
A/Residues: 1-1255 <KIN>

A/Cross-references: UNIPROT:P36332; GB:L04653; NID:G290609; PIDN:AC19319.1; PID:G290611
C/Superfamily: togavirus structural polypeptide

C/Keywords: coat protein; glycoprotein; polypeptide; transmembrane protein
F:1-275/Product: coat protein #status predicted <CTP>

F:276-334/Product: membrane glycoprotein E3 #status predicted <MG3>
F:335-757/Product: membrane glycoprotein E2 #status predicted <MG2>

F:702-722/Domain: transmembrane #status predicted <TM1>
F:758-813/Product: 6K protein #status predicted <K6>

F:795-814/Domain: transmembrane #status predicted <TM2>
F:814-1255/Product: membrane glycoprotein E1 #status predicted <MG1>

F:1332-1249/Domain: transmembrane #status predicted <TM3>
F:47,266,652,947/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 48.5%; Score 48; DB 1; Length 1255;
Best Local Similarity 50.0%; Pred. No. 21;

Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SSDGLMNNQOTLFL 16

Db 763 SLDLHNNNQOMFWIQ 778

RESULT 9

D44213

structural polypeptide - Venezuelan equine encephalitis virus (strain 3880)
C:Species: Venezuelan equine encephalitis virus
C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C/Accession: D44213

R:Kinney, R.M.; Teuchiya, K.R.; Snider, J.M.; Trent, D.W.
Virology 191, 569-580, 1992

A>Title: Genetic evidence that epizootic Venezuelan equine encephalitis (VEE) viruses ma
A/Reference number: A44213; MUID:93079859; PMID:1448915
A/Accession type: genomic RNA

A/Molecule type: genomic RNA
A/Residues: 1-1255 <KIN>

A/Cross-references: UNIPROT:P36332; GB:L00930; NID:G323706; PIDN:AC19325.1; PID:G137509
C/Superfamily: togavirus structural polypeptide

C/Keywords: coat protein; glycoprotein; polypeptide; transmembrane protein
F:1-275/Product: coat protein #status predicted <CTP>

F:276-334/Product: membrane glycoprotein E3 #status predicted <MG3>
F:335-757/Product: membrane glycoprotein E2 #status predicted <MG2>

F:702-722/Domain: transmembrane #status predicted <TM1>
F:758-813/Product: 6K protein #status predicted <K6>

F:795-814/Domain: transmembrane #status predicted <TM2>
F:814-1255/Product: membrane glycoprotein E1 #status predicted <MG1>

F:1332-1249/Domain: transmembrane #status predicted <TM3>
F:47,266,652,947/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 48.5%; Score 48; DB 1; Length 1255;
Best Local Similarity 50.0%; Pred. No. 21;

Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SSDGLMNNQOTLFL 16

Db 763 SLDLHNNNQOMFWIQ 778

RESULT 10

T06663

DNA (cytosine-5)-methyltransferase (EC 2.1.1.37) T6G15.160 - Arabidopsis thaliana
N/Alternate names: protein T6G15.160
C/Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C/Accession: T06663

R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X
submitted to the Protein Sequence Database, April 1999

A/Reference number: 215791
A/Accession: T06663

A/Molecule type: DNA
A/Residues: 1-1404 <BEV>

A/Cross-references: UNIPROT:Q9T011; EMBL:AL049656; GSPDB:GN00062; ATSP:T6G15.160
A/Experimental source: cultivar Columbia; BAC clone T6G15

C/Genetics:
A/Gene: ATSP:T6G15.160

A/Map position: 4
A/Intons: 57/1; 348/3; 371/2; 872/3; 927/3; 989/1; 1026/2; 1086/3; 1152/3; 1245/3; 1288/

C/Keywords: methyltransferase; S-adenosylmethionine

Query Match 47.5%; Score 47; DB 2; Length 1404;
Best Local Similarity 57.1%; Pred. No. 35;

Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 SDGLMNNQOTLFL 15

Db 1095 SDGSMSKNQOML 1108

RESULT 11

S46008

probable carboxypeptidase C (EC 3.4.16.5) YBR139w - yeast (Saccharomyces cerevisiae)
N/Alternate names: hypothetical protein YBR1015, hypothetical protein YBR139w
C/Species: Saccharomyces cerevisiae
C>Date: 26-Aug-1994 #sequence_revision 05-Jan-1996 #text_change 16-Aug-2004
C/Accession: S46008; S46581

R:Beck, A.M.; Herbert, C.J.; Nasr, F.; Slonimski, P.P.; Zagulski, M.
submitted to the Protein Sequence Database, August 1994

A/Reference number: S45995
A/Accession: S46008

A/Molecule type: DNA
A/Residues: 1-508 <BEC>

A/Cross-references: UNIPROT:P38109; EMBL:Z36008; NID:G536435; PIDN:CA65097.1; PID:G53644
A/Experimental source: strain S288C

R:Beck, A.M.; Cullin, C.; Grzybowski, E.; Lacroite, F.; Nasr, F.; Ozier-Kalogeropoulos,
Yeast 10(Suppl.A), S1-S11, 1994

A>Title: The sequence of 29,7kb from the right arm of chromosome II reveals 13 complete c
A/Reference number: S46569; MUID:94378717; PMID:8091856

A/Accession type: DNA
A/Residues: 1-508 <BE2>

A/Cross-references: EMBL:X75891; NID:G496856; PIDN:CA53497.1; PID:G496869
A/Experimental source: strain S288C

C/Genetics:
A/Cross-references: SGD:S0000343

A/Map position: 2R
A/Note: MIPS:YBR139w

C/Superfamily: Serine carboxypeptidase
C/Keywords: hydrolase, serine carboxypeptidase; transmembrane protein

F:6-22/Domain: transmembrane #status predicted <TM1>
F:219,415,474/Active site: Ser Asp, His #status predicted

Query Match 46.5%; Score 46; DB 1; Length 508;
Best Local Similarity 63.6%; Pred. No. 16;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 6 WNNNQOTLFL 16

Db 160 WNNNASMIFLE 170

RESULT 12

```

A:Molecule type: protein
A:Residues: 201-259, 'C', 261-281, 'ND', 284-301, <DRI>
C:Comment: Calcium is required for ligand binding.
C:Superfamily: hepatic lectin; C-type lectin homology
C:Keywords: endocytosis; glycoprotein; lectin; liver;
F:2-60/Domain: intracellular #status predicted <INT>
F:61-77/Domain: transmembrane #status predicted <TRN>
F:78-301/Domain: extracellular #status predicted <EXT>
F:170-293/Domain: C-type lectin homology <LCH>
F:97,119,165/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match          45.5%; Score 45; DB 1; Length 301;
Best Local Similarity 72.7%; Pred. No. 13;
Matches      8; Conservative      1; Mismatches      2; Indels      0; Gaps      0;

QY      2 SDGLMNNNQTO 12
         |||||:|:|:|
Db       276 SDGLMNDNFCQ 286

RESULT 14
S29159
glutamate receptor, NMDA-sensitive, epsilon-1 chain precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S29159
R:Meguro, H.; Mori, H.; Araki, K.; Kushiya, E.; Kutsuwada, T.; Yamazaki, M.; Kumaiishi,
Nature 357, 70-74, 1992
A:Title: Functional characterization of a heteromeric NMDA receptor channel expressed fr
A:Reference number: S29159; MUID:92244361; PMID:1374164
A:Accession: S29159
A:Molecule type: mRNA
A:Residues: 1-1464 <MEG>
A:Cross-references: UNIPROT:P35436; EMBL:DI0217; NID:9220410; PIDN:BA01069.1; PID:92204
C:Superfamily: N-methyl-D-aspartate receptor 2A; glutamate receptor homology
C:Keywords: ion channel; neurotransmitter receptor; transmembrane protein
F:1-19/Domain: signal sequence #status predicted <Sig>
F:20-1464/Domain: glutamate receptor epsilon-1 chain #status predicted <MAT>
F:1428-854/Domain: glutamate receptor homology <GRN>
F:557-576/Domain: transmembrane #status predicted <TM1>
F:599-618/Domain: transmembrane #status predicted <TM2>
F:629-647/Domain: transmembrane #status predicted <TM3>
F:817-837/Domain: transmembrane #status predicted <TM4>

Query Match          45.5%; Score 45; DB 1; Length 1464;
Best Local Similarity 47.1%; Pred. No. 76;
Matches      8; Conservative      5; Mismatches      4; Indels      0; Gaps      0;

QY      1 SSDGLMNNNQTOFLFH 17
         :|||:|:|:|:|
Db      1174 NEDGLPNNDQYKLYAKH 1190

RESULT 15
A:Molecule type: DNA
A:Residues: 1-1464 <MON>
A:Cross-references: UNIPROT:Q00959; GB:M91561; NID:92905605; PIDN:AC03565.1; PID:929058
A:Experimental source: brain
A>Note: Sequence modified after extraction from NCBI backbone
R:Inohi, T.; Moriyoehi, K.; Sugihara, H.; Sakurada, K.; Kadotani, H.; Yokoi, M.; Akazawa
J. Biol. Chem. 268, 2835-2843, 1993
A:Title: Molecular characterization of the family of the N-methyl-D-aspartate receptor s
A:Reference number: A45219; MUID:931515102; PMID:8428958

```

A/Accession: A45219
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-245, 'P', 247-757, 'T', 759-1464 <ISH>
A/Cross-references: GB:D13211; NID:G286233; PIDN:BAA02498.1; PID:G286234
A/Experimental source: Brain
A/Note: Sequence extracted from NCBI backbone (NCBIP:124262)
C/Superfamily: N-methyl-D-aspartate receptor 2A; glutamate receptor homology
F:428-854/Domain: glutamate receptor homology <GRH>

Query Match 45.5%; Score 45; DB 2; Length 1464;
Best Local Similarity 47.1%; Pred. No. 76;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 SSDGLMNNNOTQFLFHS 17
Db 1174 NEDGLPNNDOYKYLAKH 1190

RESULT 16
A29476
muscarinic acetylcholine receptor M4 - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
C/Accession: A29476
R/Braun, T.; Schofield, P.R.; Shivers, B.D.; Pritchett, D.B.; Seeburg, P.H.
Biochem. Biophys. Res. Commun. 149, 125-132, 1987
A/Title: A novel subtype of muscarinic receptor identified by homology screening.
A/Reference number: A29476; MUID:88077068; PMID:3120722
A/Accession: A29476
A/Molecule type: mRNA
A/Residues: 1-589 <BRA>
A/Cross-references: UNIPROT:P08483; GB:M18088; NID:G202657; PIDN:AAA40659.1; PID:G202658
A/Experimental source: brain
C/Superfamily: vertebrate rhodopsin
C/Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; transmembrane
F:67-90/Domain: transmembrane #status predicted <TM1>
F:104-124/Domain: transmembrane #status predicted <TM2>
F:142-163/Domain: transmembrane #status predicted <TM3>
F:184-206/Domain: transmembrane #status predicted <TM4>
F:230-251/Domain: transmembrane #status predicted <TM5>
F:492-512/Domain: transmembrane #status predicted <TM6>
F:527-545/Domain: transmembrane #status predicted <TM7>

Query Match 44.4%; Score 44; DB 2; Length 589;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 1 SSDGLMNNNOTQFLFHS 18
Db 331 SSSDSWNNNDAAASLENS 348

RESULT 17
B29514
muscarinic acetylcholine receptor M3 - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
C/Accession: B94518; B94293; B37121; B29514
R/Bonner, T.I.
submitted to GenBank, July 1987
A/Reference number: A94518
A/Accession: B94518
A/Molecule type: mRNA
A/Residues: 1-589 <BO1>
A/Cross-references: UNIPROT:P08483
R/Bonner, T.I.; Buckley, N.V.; Young, A.C.; Brann, M.R.
Science 237, 527-532, 1987
A/Title: Identification of a family of muscarinic acetylcholine receptor genes.
A/Reference number: A94293; MUID:87263421; PMID:3037705
A/Accession: B94293
A/Molecule type: mRNA
A/Residues: 1-269; 463-589 <BO2>

A/Experimental source: cerebral cortex
A/Note: only a part of the protein translation is given; none of the nucleotide sequence
R/Kurtenbach, E.; Curtiss, C.A.M.; Pedder, E.K.; Altken, A.; Harris, A.C.M.; Hulme, E.C.
J. Biol. Chem. 265, 13702-13709, 1990
A/Title: Muscarinic acetylcholine receptors. Peptide sequencing identifies residues inno)
A/Reference number: A37121; MUID:90337982; PMID:2380182
A/Accession: B37121
A/Status: preliminary
A/Molecule type: protein
A/Residues: 104-166 <KUR>
C/Superfamily: vertebrate rhodopsin
C/Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; transmembrane
F:67-90/Domain: transmembrane #status predicted <TM1>
F:104-124/Domain: transmembrane #status predicted <TM2>
F:142-163/Domain: transmembrane #status predicted <TM3>
F:184-206/Domain: transmembrane #status predicted <TM4>
F:230-251/Domain: transmembrane #status predicted <TM5>
F:492-512/Domain: transmembrane #status predicted <TM6>
F:527-545/Domain: transmembrane #status predicted <TM7>
F:6,15,41,48,52/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 44.4%; Score 44; DB 2; Length 589;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 1 SSDGLMNNNOTQFLFHS 18
Db 331 SSSDSWNNNDAAASLENS 348

RESULT 18
S10128
muscarinic acetylcholine receptor M4 - human
C/Species: Homo sapiens (man)
C/Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C/Accession: S10128
R/Peralta, E.G.; Ashkenazi, A.; Winslow, J.W.; Smith, D.H.; Ramchandran, J.; Capon, D.J.
EMBO J. 6, 3923-3929, 1987
A/Title: Distinct primary structures, ligand-binding properties and tissue-specific expression of human muscarinic acetylcholine receptors.
A/Reference number: S04326; MUID:88166633; PMID:3443095
A/Accession: S10128
A/Molecule type: DNA
A/Residues: 1-590 <PER>
A/Cross-references: UNIPROT:P20309; EMBL:X15266; NID:G32323; PIDN:CAA3337.1; PID:G32324
C/Superfamily: vertebrate rhodopsin
C/Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; transmembrane
F:68-95/Domain: transmembrane #status predicted <TM1>
F:105-131/Domain: transmembrane #status predicted <TM2>
F:143-164/Domain: transmembrane #status predicted <TM3>
F:185-207/Domain: transmembrane #status predicted <TM4>
F:231-252/Domain: transmembrane #status predicted <TM5>
F:493-513/Domain: transmembrane #status predicted <TM6>
F:525-546/Domain: transmembrane #status predicted <TM7>
F:5,6,15,41/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 44.4%; Score 44; DB 2; Length 590;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 1 SSDGLMNNNOTQFLFHS 18
Db 332 SSSDSWNNNDAAASLENS 349

RESULT 19
S01114
muscarinic acetylcholine receptor M2, glandular - pig
A/Alternate names: muscarinic acetylcholine receptor III
C/Species: Sus scrofa domestica (domestic pig)
C/Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 09-Jul-2004
C/Accession: S01114
R/Akiba, I.; Kubo, T.; Maeda, A.; Bujo, H.; Nakai, U.; Mishina, M.; Numa, S.
FEBS Lett. 235, 257-261, 1988

A>Title: Primary structure of porcine muscarinic acetylcholine receptor III and antagonist
A:Reference number: S01114; MUID:88296835; PMID:3402600
A:Accession: S01114
A:Molecule type: DNA
A:Residues: 1-590 <AKI>
A:Cross-references: UNIPROT:P11483; EMBL:X12712; NID:g1861; PIDN:CAA11215.1; PID:g1862
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; transmembrane
F:68-91/Domain: transmembrane #status predicted <TM1>
F:105-125/Domain: transmembrane #status predicted <TM2>
F:143-164/Domain: transmembrane #status predicted <TM3>
F:185-207/Domain: transmembrane #status predicted <TM4>
F:231-252/Domain: transmembrane #status predicted <TM5>
F:493-513/Domain: transmembrane #status predicted <TM6>
F:528-546/Domain: transmembrane #status predicted <TM7>

Query Match 44.4%; Score 44; DB 2; Length 590;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

OY 1 SSDGLMNNQOTLFLEHS 18
Db 332 SSSDSWNNNDAAASLENS 349

RESULT 20
S47572
muscarinic acetylcholine receptor m3 - bovine
C:Species: Bos primigenius taurus (cattle)
C>Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 09-Jul-2004
C:Accession: S47572
R:Lee, P.H.K.; Hodges, P.K.; Glickman, F.; Chang, K.J.
Biochim. Biophys. Acta 1223, 151-154, 1994
A>Title: Cloning and expression of a cDNA encoding bovine muscarinic acetylcholine m3 receptor
A:Reference number: S47572; MUID:94339178; PMID:8061048
A:Accession: S47572
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-590 <LEE>
A:Cross-references: UNIPROT:P41984; EMBL:U08286; NID:9520465; PIDN:AAAS1866.1; PID:95204
C:Superfamily: vertebrate rhodopsin
C:Keywords: neurotransmitter receptor

Query Match 44.4%; Score 44; DB 2; Length 590;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

OY 1 SSDGLMNNQOTLFLEHS 18
Db 332 SSSDSWNNNDAAASLENS 349

RESULT 21
A55019
muscarinic acetylcholine receptor, M3 isoform - chicken
C:Species: Gallus gallus (chicken)
C>Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 09-Jul-2004
C:Accession: A55019
R:Gadbut, A.P.; Galper, J.B.
J. Biol. Chem. 269, 25823-25829, 1994
A>Title: A novel M-3 muscarinic acetylcholine receptor is expressed in chick atrium and
A:Reference number: A55019; MUID:95014393; PMID:7929287
A:Accession: A55019
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-639 <GAD>
A:Cross-references: UNIPROT:P49578; GB:L10617; NID:9530097; PIDN:AA65961.1; PID:9530098
C:Superfamily: vertebrate rhodopsin
C:Keywords: neurotransmitter receptor

Query Match 44.4%; Score 44; DB 2; Length 639;
Best Local Similarity 50.0%; Pred. No. 44;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

OY 1 SSDGLMNNQOTLFLEHS 18
Db 380 SSSDSWNNNDAAASLENS 397

RESULT 22
T35238
probable secreted cellulase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C:Accession: T35238
R:Seeger, K.J.; Harris, D.; Parkhill, J.; Barrall, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1998
A:Reference number: 221572
A:Accession: T35238
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-973 <SEE>
A:Cross-references: UNIPROT:O86728; EMBL:AL031515; PIDN:CAA20643.1; GSPDB:GN00070; SCOEPD
A:Experimental source: strain A3(2)
A:Gene: SCOEDB:SC5C7.31c

Query Match 44.4%; Score 44; DB 2; Length 973;
Best Local Similarity 63.6%; Pred. No. 70;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 3 DGLMNNQOTOL 13
Db 872 DGWENNQDAL 882

RESULT 23
G8EBC
beta-galactosidase (EC 3.2.1.23) alpha chain - Escherichia coli (strain K-12)
N:Alternate names: beta-D-galactoside galactohydrolase; lactase; phospho-beta-D-galactose
C:Species: Escherichia coli
C>Date: 30-Sep-1987 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: A65096; A25751; S09206
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A>Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: A65096
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1042 <BLAT>
A:Cross-references: UNIPROT:P06864; UNIPROT:Q47170; GB:AE000389; GB:U00096; NID:g1789451
A:Experimental source: strain K-12, substrain W61655
R:Stokes, H.W.; Betts, P.W.; Hall, B.G.
Mol. Biol. Evol. 2, 469-477, 1985
A>Title: Sequence of the ebgA gene of Escherichia coli: comparison with the lacZ gene.
A:Reference number: A93056; MUID:88216133; PMID:33939707
A:Accession: A25751
A:Molecule type: DNA
A:Residues: 80-476, 'R', 478-651, 'S', 653-660, 'P', 662-674, 'CRSWTPARKS', 685-724, 'KCA5V', 730-
A:Cross-references: GB:X03228; GB:M13700; GB:M13796; NID:941311; PIDN:CAA26978.1; PID:94
J.Fowler, A.V.; Smith, P.J.
J. Biol. Chem. 258, 10204-10207, 1983
A>Title: The active site regions of lacZ and ebg beta-galactosidases are homologous.
A:Reference number: A92390; MUID:83290932; PMID:6411710
A:Contents: annotation; active site regions
R:Hall, B.G.; Betts, P.W.; Woollam, J.C.
Gene 123, 635-648, 1989
A>Title: DNA sequence analysis of artificially evolved ebg enzyme and ebg repressor gene
A:Reference number: S09205; MUID:90128210; PMID:2515108
A:Accession: S09206
A:Molecule type: DNA
A:Residues: 13-476, 'R', 478-651, 'S', 653-660, 'P', 662-779, 'W', 780-1037, 'T', 1039-1042 <HAL>
A:Cross-references: EMBL:X5031; NID:941307; PIDN:CAA36274.1; PID:941309
C:Comment: The wild-type enzyme is an ineffective lactase. Two classes of point mutation:

C/Genetics:
A:Gene: ebga
A:Map position: 68 min
C/Complex: homohexamer
C/Superfamily: beta-galactosidase
C/Keywords: glycosidase; hexamer; hydrolase; magnesium
F:413,415,461/Binding site: magnesium (Glu, His, Glu) \$status predicted
F:461,502,524/Active site: Glu, Tyr, Glu \$status predicted

Query Match 44.4%; Score 44; DB 1; Length 1042;
Best Local Similarity 46.7%; Pred. No. 76;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 3 DGLMNNQTOLEH 17
||| | | |
Db 788 EGLWQPNHLQIMQH 802

RESULT 24
E85968
evolved beta-D-galactosidase alpha subunit [imported] - Escherichia coli (strain O157:H7)
C/Species: Escherichia coli
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C/Accession: E85968
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A/Reference number: A85480; MUID:21074935; PMID:11206551
A/Accession: E85968
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1042 <HAY>
A/Cross-references: UNIPROT:O8XAM9; GB:AB005174; NID:g12517660; PIDN:AA68209.1; GSPDB:G
C/Genetics:
A:Gene: ebga
C/Superfamily: beta-galactosidase

Query Match 44.4%; Score 44; DB 2; Length 1042;
Best Local Similarity 46.7%; Pred. No. 76;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 3 DGLMNNQTOLEH 17
||| | | |
Db 788 EGLWQPNHLQIMQH 802

RESULT 25
F91123
evolved beta-D-galactosidase alpha subunit [imported] - Escherichia coli (strain O157:H7)
C/Species: Escherichia coli
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C/Accession: F91123
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A/Reference number: A99629; MUID:21156231; PMID:11258796
A/Accession: F91123
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1042 <HAY>
A/Cross-references: UNIPROT:O8XAM9; GB:BA000007; PIDN:BA37381.1; PID:g13363431; GSPDB:G
C/Genetics:
A:Gene: EC83958
C/Superfamily: beta-galactosidase

Query Match 44.4%; Score 44; DB 2; Length 1042;
Best Local Similarity 46.7%; Pred. No. 76;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 3 DGLMNNQTOLEH 17
||| | | |
Db 788 EGLWQPNHLQIMQH 802

RESULT 26
C82931
hypothetical protein Uui26 [imported] - Ureaplasma urealyticum
C/Species: Ureaplasma urealyticum
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C/Accession: C82931
R:Glaser, J.L.; Leikowitz, E.J.; Glaser, J.S.; Heiner, C.R.; Chen, E.Y.; Caswell, G.H.
submitted to GenBank, February 2000
A/Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
A/Reference number: A82870
A/Accession: C82931
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1084 <GLA>
A/Cross-references: GB:AF222894; NID:g6899078; PIDN:AAF30532.1; GSPDB:GN001;
A/Experimental source: serovar 3, biovar 1
C/Genetics:
A:Gene: Uui26
A/Genetic code: SCC3

Query Match 44.4%; Score 44; DB 2; Length 1084;
Best Local Similarity 50.0%; Pred. No. 79;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 5 LMNNQTOLEH 18
||| | | |
Db 479 LMASNNLYLDYS 492

RESULT 27
S76938
hypothetical protein - Synechocystis sp. (strain PCC 6803)
C/Species: Synechocystis sp.
A/Variety: PCC 6803
C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 16-Aug-2004
C/Accession: S76938
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,
DNA Res. 3, 109-136, 1996
A/Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
sp.
A/Reference number: S74322; MUID:97061201; PMID:8905231
A/Accession: S76938
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-254 <KAN>
A/Cross-references: UNIPROT:P74730; EMBL:D90917; GB:AB001339; NID:g1653836; PIDN:BA11885C
C/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C/Superfamily: ATP-binding cassette homology
C/Keywords: ATP; nucleotide binding; P-loop
F:45-225/Domain: ATP-binding cassette homology <ABC>
F:62-69/Region: nucleotide-binding motif A (P-loop)

Query Match 43.4%; Score 43; DB 2; Length 254;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 3 DGLMNNQTOLEH 18
||| | | |
Db 193 DKLGHHTTILVSHS 208

RESULT 28
AG2598
conserved hypothetical protein phng [imported] - Agrobacterium tumefaciens (strain C58, I
C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 12-Jul-2004
C/Accession: AG2598

Riwood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Karp, P.; Gillet, W.; Grant, C.; Guenther, D.; Kuyavirin, T.; Levy, R.; Li, M.; McChell, Science 294, 2317-2323, 2001

A:Author: Yoo, H.; Rao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E.W.

A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.

A:Reference number: AB2577; MUID:21608550; PMID:11743193

A:Accession: AG2598

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-153 <KUR>

A:Cross-references: UNIPROT:Q8U1W0; GB:AE008688; PIDN:AAL41205.1; PID:g17738507; GSPDB:G

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: pnhg

A:Map position: circular chromosome

C:Superfamily: carbon-phosphorus lyase, component pnhg

Query Match 42.4%; Score 42; DB 2; Length 153;

Best Local Similarity 50.0%; Pred. No. 19;

Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 DGLMNNQTLFLE 16

DB 103 DALMQESQTRGRIE 116

RESULT 29

pnhg protein [imported] - *Agrobacterium tumefaciens* (strain C58, Cereon)

C:Species: *Agrobacterium tumefaciens*

C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 12-Jul-2004

C:Accession: H97380

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tumefaciens*

A:Reference number: A97359; MUID:21608551; PMID:11743194

A:Accession: H97380

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-153 <KUR>

A:Cross-references: UNIPROT:Q8U1W0; GB:AE007869; PIDN:AAK66001.1; PID:g15155066; GSPDB:G

C:Genetics:

A:Gene: AGR_C_303

A:Map position: circular chromosome

C:Superfamily: carbon-phosphorus lyase, component pnhg

Query Match 42.4%; Score 42; DB 2; Length 153;

Best Local Similarity 50.0%; Pred. No. 19;

Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 DGLMNNQTLFLE 16

DB 103 DALMQESQTRGRIE 116

RESULT 30

hypothetical protein b2847 - *Escherichia coli* (strain K-12)

C:Species: *Escherichia coli*

C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004

C:Accession: H65067

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C. Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of *Escherichia coli* K-12.

A:Reference number: A64720; MUID:97426617; PMID:9276503

A:Accession: H65067

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-269 <BLAT>

A:Cross-references: UNIPROT:Q46942; GB:AE000368; GB:U00096; MUID:g2367165; PIDN:AACT5886; A:Experimental source: strain K-12, substrain MG1655

Query Match 42.4%; Score 42; DB 2; Length 269;

Best Local Similarity 58.3%; Pred. No. 35;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 LMNNNQTLFLE 16

DB 173 LMNNNQPKPFK 184

RESULT 31

probable sensory transducer [imported] - *Escherichia coli* (strain O157:H7, substrain RIM

C:Species: *Escherichia coli*

C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004

C:Accession: H91091

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and geno

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: H91091

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-269 <HAY>

A:Cross-references: UNIPROT:Q8X613; GB:BA000007; PIDN:BA837127.1; PID:g13363176; GSPDB:G

A:Experimental source: strain O157:H7, substrain RIMD 0509952

C:Genetics:

A:Gene: EC83704

Query Match 42.4%; Score 42; DB 2; Length 269;

Best Local Similarity 58.3%; Pred. No. 35;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 LMNNNQTLFLE 16

DB 173 LMNNNQPKPFK 184

RESULT 32

probable sensory transducer ygeI [imported] - *Escherichia coli* (strain O157:H7, substrain

C:Species: *Escherichia coli*

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004

C:Accession: C85937

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew iller, L.; Grobeck, B.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamouls, K.; Apodaca, Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: C85937

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-269 <STO>

A:Cross-references: UNIPROT:Q8X613; GB:AE005174; MUID:g12517339; PIDN:AG57959.1; GSPDB:G

A:Experimental source: strain O157:H7, substrain EDU933

C:Genetics:

A:Gene: ygeI

Query Match 42.4%; Score 42; DB 2; Length 269;

Best Local Similarity 58.3%; Pred. No. 35;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 LMNNNQTLFLE 16

DB 173 LMNNNQPKPFK 184

RESULT 33

hypothetical protein 10 - red alga (*Porphyra purpurea*) chloroplast

C/Species: chloroplast Porphyra purpurea
C/Date: 19-Mar-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004
C/Accession: S73153
R/Reith, M.; Munholland, J.
Plant Mol. Biol. Rep. 13, 333-335, 1995
A/Title: Complete nucleotide sequence of the Porphyra purpurea chloroplast genome.
A/Reference number: S73108
A/Accession: S73153
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-278 <Ref>
A/Cross-references: UNIPROT:P51232; EMBL:U38804; NID:g1276652; PID:g1276698
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995
C/Genetics:
A/Gene: ycf10
A/Genome: chloroplast
C/Superfamily: maize chloroplast protein cema
C/Keywords: chloroplast

Query Match 42.4%; Score 42; DB 2; Length 278;
Best Local Similarity 43.8%; Pred. No. 36;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Oy 3 SDGLMNNQQLFLEHS 18
||| |::| |
Db 85 DYLWNOEQPKIFINSS 100

RESULT 34

VGX837
glycoprotein VP7 precursor - human rotavirus A (serotype 1 strain M37)
N/Alternate names: outer capsid protein VP7

C/Species: human rotavirus A
C/Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
C/Accession: C27620

R/Green, K.Y.; Mitchum, K.; Gorziglia, M.; Hoshino, Y.; Kapikian, A.Z.; Chanock, R.M.; F
Virology 161, 153-159, 1987
A/Title: Comparison of the amino acid sequences of the major neutralization protein of R

A/Reference number: A27620; MUID:88044489; PMID:2823458
A/Accession: C27620

A/Molecule type: genomic RNA
A/Residues: 1-326 <GR>

A/Cross-references: UNIPROT:P11852
C/Genetics:

A/Map position: segment 9
C/Superfamily: rotavirus glycoprotein VP7

C/Keywords: coat protein; glycoprotein; transmembrane protein
F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-326/Product: glycoprotein VP7 #status predicted <GPV>
F:32-48/Region: hydrophobic #status predicted

F:69,238/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 42.4%; Score 42; DB 1; Length 326;
Best Local Similarity 50.0%; Pred. No. 43;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Oy 2 SDGLMNNQQLFLEHS 15
||| |::| |
Db 94 SDGEMKSLSQMFL 107

RESULT 35

A97535
ispD/ISP bifunctional enzyme [imported] - Agrobacterium tumefaciens (strain C58, Cereot

C/Species: Agrobacterium tumefaciens
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C/Accession: A97535

R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001

A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A/Reference number: A97535; MUID:21608551; PMID:11743194
A/Accession: A97535

A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-400 <Kur>
A/Cross-references: UNIPROT:Q8UF4; GB:AE07869; PID:AAK87234.1; PID:g15156518; GSPDB:G
C/Genetics:
A/Gene: AGR_C_2659
A/Map position: circular chromosome
C/Superfamily: bifunctional 4-diphosphocytidylyl-2-methyl-D-erythritol synthase/2C-methyl-1

Query Match 42.4%; Score 42; DB 2; Length 400;
Best Local Similarity 41.2%; Pred. No. 54;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Oy 2 SDGLMNNQQLFLEHS 18
||| |::| |
Db 303 SDPQWKGASRIFFLEHA 319

RESULT 36

AC2754
ISP/ISP bifunctional enzyme [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C/Accession: AC2754

R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001

A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E
ster, E.W.

A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193

A/Accession: AC2754
A/Status: preliminary

A/Molecule type: DNA
A/Residues: 1-400 <Kur>

A/Cross-references: UNIPROT:Q8UF4; GB:AE008688; PID:AAI42449.1; PID:g17739863; GSPDB:G
A/Experimental source: strain C58 (Dupont)

C/Genetics:

A/Gene: ispF
A/Map position: circular chromosome

C/Superfamily: bifunctional 4-diphosphocytidylyl-2-methyl-D-erythritol synthase/2C-methyl-1

Query Match 42.4%; Score 42; DB 2; Length 400;
Best Local Similarity 41.2%; Pred. No. 54;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Oy 2 SDGLMNNQQLFLEHS 18
||| |::| |
Db 303 SDPQWKGASRIFFLEHA 319

RESULT 37

T05439
hypothetical protein F7K2.30 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C/Accession: T05439

R/Beyan, M.; Medler, H.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Schueller;
submitted to the Protein Sequence Database, November 1998

A/Reference number: Z15416
A/Accession: T05439

A/Molecule type: DNA
A/Residues: 1-457 <BEV>

A/Cross-references: UNIPROT:Q9SUX4; EMBL:AL033545
A/Experimental source: cultivar Columbia; BAC clone F7K2

C/Genetics:
A/Map position: 4
A/Introns: 8/3; 107/3; 309/3; 327/3
A/Note: F7K2.30

Query Match 42.4%; Score 42; DB 2; Length 457;
Best Local Similarity 50.0%; Pred. No. 63;

Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 Oy 5 LMNNQQLFLE 16
 :||:|:|:
 Db 11 VMNDQTRFYLE 22

RESULT 38

A58932
 cytochrome C-type biogenesis protein CCMF - Cyanidioschyzon merolae mitochondrion
 C/Species: mitochondrion Cyanidioschyzon merolae
 C/Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
 C/Accession: A58932
 R/Ohta, N.; Sato, N.; Kuroiwa, T.
 Nucleic Acids Res. 26, 5190-5198, 1998
 A/Title: Structure and organization of the mitochondrial genome of the unicellular red alga
 A/Reference number: A58930; MUID:99030526; PMID:9801318
 A/Accession: A58932
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-620 <ARN>
 A/Cross-references: UNIPROT:Q9Z2P7; GB:D89861; NID:G4115781; PIDN:BA036527.1; PID:G41157
 C/Genetics:
 A/Gene: yefR; ccmF
 A/Genome: mitochondrion
 C/Superfamily: nfeF protein
 C/Keywords: mitochondrion

Query Match 42.4%; Score 42; DB 2; Length 620;
 Best Local Similarity 63.6%; Pred. No. 88;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 5 LMNNQQLFLE 15
 :||:|:|:
 Db 91 LMSNNEGLFL 101

RESULT 39

S75724
 hypothetical protein slr0907 - Synecchocystis sp. (strain PCC 6803)
 C/Species: Synecchocystis sp.
 A/Variety: PCC 6803
 C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
 C/Accession: S75724
 R.; Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
 O., K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
 DNA Res. 3, 109-136, 1996
 A/Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecchocystis
 sp.
 A/Reference number: S74322; MUID:97061201; PMID:8905231
 A/Accession: S75724
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-1014 <KAN>
 A/Cross-references: UNIPROT:O55174; EMBL:D64003; GB:AB001339; NID:G1001200; PIDN:BA01045
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 42.4%; Score 42; DB 2; Length 1014;
 Best Local Similarity 43.8%; Pred. No. 1.5e+02;
 Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Oy 1 SSDGLMNNQQLFLE 16
 :||:|:|:
 Db 819 ASHGIMSHSTASQLPE 834

RESULT 40

S72645
 pilY protein - Pseudomonas aeruginosa
 C/Species: Pseudomonas aeruginosa
 C/Date: 29-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 09-Jul-2004
 C/Accession: S72645; S72633
 R.; Alm, R.A.; Hallinan, J.P.; Watson, A.A.; Mattick, J.S.

submitted to the EMBL Data Library, April 1996
 A/Description: Fimbrial biogenesis genes of Pseudomonas aeruginosa: pilW and pilX increas
 omologue:
 A/Reference number: S72643
 A/Accession: S72645
 A/Molecule type: DNA
 A/Residues: 1-1161 <ALM>

A/Cross-references: UNIPROT:O51536; EMBL:L76605; NID:G1246298; PIDN:AA093502.1; PID:G12462
 A/Experimental source: strain PAOI
 R.; Alm, R.A.; Hallinan, J.P.; Watson, A.A.; Mattick, J.S.
 Mol. Microbiol. 22, 161-173, 1996
 A/Title: Fimbrial biogenesis genes of Pseudomonas aeruginosa: pilW and pilX increase the

A/Reference number: S72631; MUID:97055431; PMID:8899718
 A/Accession: S72633
 A/Status: nucleic acid sequence not shown
 A/Molecule type: DNA
 A/Residues: 657-913 <ALW>
 A/Cross-references: EMBL:L76605
 A/Note: only a part of the translation is shown
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1996
 C/Genetics:
 A/Gene: pilY1
 C/Function:
 A/Description: may be involved in assembly of type 4 fimbriae
 C/Keywords: fimbria; membrane protein

Query Match 42.4%; Score 42; DB 2; Length 1161;
 Best Local Similarity 50.0%; Pred. No. 1.8e+02;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Oy 2 SSDGLMNNQQL 13
 :||:|:|:
 Db 357 TDGLMNNDSANV 368

Search completed: December 30, 2004, 16:13:57
 Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 16:23:09 : Search time 37 Seconds
(without alignments)
32.263 Million cell updates/sec

Title: US-10-718-321-1

Perfect score: 99
Sequence: 1 SSDGLMNNQTOLEHS 18

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 187706

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 150 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34.3	19	4	US-09-441-502B-68	Sequence 68, Appl
2	32.3	12	3	US-08-817-177-2	Sequence 2, Appl
3	32.3	12	5	PCT-US95-12686-2	Sequence 2, Appl
4	32.3	19	1	US-08-218-025A-166	Sequence 166, App
5	31.3	7	1	US-08-277-660A-22	Sequence 22, Appl
6	31.3	7	1	US-08-424-957-7	Sequence 7, Appl
7	31.3	7	3	US-09-035-686-7	Sequence 6, Appl
8	30.3	13	3	US-08-593-500-6	Sequence 6, Appl
9	30.3	13	3	US-08-195-006-6	Sequence 6, Appl
10	30.3	13	5	PCT-US94-07644A-6	Sequence 25, Appl
11	30.3	14	4	US-09-638-203-25	Sequence 82, Appl
12	30.3	20	1	US-08-218-025A-82	Sequence 82, Appl
13	30.3	20	3	US-08-612-973-73	Sequence 82, Appl
14	30.3	20	3	US-08-612-973-82	Sequence 82, Appl
15	30.3	20	3	US-08-927-597-73	Sequence 82, Appl
16	30.3	20	3	US-08-927-597-82	Sequence 82, Appl
17	30.3	20	4	US-08-635-886C-38	Sequence 38, Appl
18	30.3	20	4	US-08-635-886C-47	Sequence 38, Appl
19	30.3	20	4	US-08-635-886C-56	Sequence 38, Appl
20	30.3	20	4	US-08-974-690C-38	Sequence 38, Appl
21	30.3	20	4	US-08-974-690C-47	Sequence 38, Appl
22	30.3	20	4	US-08-974-690C-56	Sequence 38, Appl
23	30.3	20	4	US-08-974-690C-56	Sequence 38, Appl
24	30.3	20	4	US-08-974-690C-56	Sequence 38, Appl
25	30.3	20	4	US-08-974-690C-56	Sequence 38, Appl
26	30.3	20	4	US-08-974-690C-56	Sequence 38, Appl
27	29.3	5	1	US-08-014-979-117	Sequence 117, Appl

28	29.3	7	2	US-08-292-968-43	Sequence 43, Appl
29	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
30	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
31	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
32	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
33	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
34	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
35	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
36	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
37	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
38	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
39	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
40	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
41	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
42	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
43	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
44	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
45	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
46	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
47	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
48	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
49	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
50	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
51	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
52	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
53	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
54	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
55	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
56	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
57	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
58	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
59	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
60	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
61	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
62	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
63	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
64	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
65	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
66	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
67	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
68	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
69	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
70	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
71	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
72	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
73	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
74	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
75	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
76	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
77	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
78	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
79	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
80	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
81	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
82	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
83	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
84	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
85	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
86	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
87	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
88	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
89	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
90	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
91	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
92	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
93	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
94	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
95	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
96	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
97	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
98	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
99	29.3	7	2	US-08-467-974-43	Sequence 43, Appl
100	29.3	7	2	US-08-467-974-43	Sequence 43, Appl

101 27 27.3 10 2 US-08-679-635A-17 Sequence 17, App1
102 27 27.3 10 3 US-09-419-163-17 Sequence 17, App1
103 27 27.3 10 4 US-09-428-082B-667 Sequence 667, App
104 27 27.3 12 1 US-08-241-054-42 Sequence 42, App1
105 27 27.3 12 1 US-08-390-156A-18 Sequence 18, App1
106 27 27.3 12 1 US-08-439-817-22 Sequence 22, App1
107 27 27.3 12 1 US-08-485-508-42 Sequence 42, App1
108 27 27.3 13 1 US-08-218-025A-172 Sequence 172, App
109 27 27.3 15 1 US-08-218-025A-24 Sequence 24, App
110 27 27.3 15 4 US-09-514-245-124 Sequence 124, App
111 27 27.3 15 4 US-09-514-245-125 Sequence 125, App
112 27 27.3 15 4 US-09-514-245-150 Sequence 150, App
113 27 27.3 15 4 US-09-514-245-151 Sequence 151, App
114 27 27.3 15 5 PCT-US93-11703-36 Sequence 36, App1
115 27 27.3 15 5 PCT-US93-11703-37 Sequence 37, App1
116 27 27.3 16 3 US-08-847-844A-36 Sequence 36, App1
117 27 27.3 17 1 US-08-218-025A-153 Sequence 153, App
118 27 27.3 17 2 US-08-878-546-6 Sequence 6, App1
119 27 27.3 17 3 US-09-399-494-6 Sequence 6, App1
120 27 27.3 17 6 5266328-11 Patent No. 5266328
121 27 27.3 19 3 US-09-227-357-325 Sequence 325, App
122 27 27.3 19 4 US-09-082-358B-56 Sequence 56, App1
123 27 27.3 20 1 US-08-218-025A-37 Sequence 37, App1
124 27 27.3 20 1 US-08-614-935-11 Sequence 11, App1
125 27 27.3 20 3 US-09-130-287-11 Sequence 11, App1
126 27 27.3 6 4 US-09-170-769A-13 Sequence 13, App1
127 27 27.3 7 1 US-08-486-721A-8 Sequence 8, App1
128 27 27.3 8 3 US-08-444-818-598 Sequence 598, App
129 27 27.3 9 3 US-08-727-308-7 Sequence 7, App1
130 27 27.3 10 1 US-08-214-650-44 Sequence 44, App1
131 27 27.3 11 2 US-08-482-228-44 Sequence 44, App1
132 27 27.3 11 3 US-08-482-528-44 Sequence 44, App1
133 27 27.3 12 1 US-08-190-788A-21 Sequence 21, App1
134 27 27.3 12 1 US-08-241-054-38 Sequence 38, App1
135 27 27.3 12 1 US-08-241-054-46 Sequence 46, App1
136 27 27.3 12 1 US-08-390-156A-16 Sequence 16, App1
137 27 27.3 12 1 US-08-390-156A-20 Sequence 20, App1
138 27 27.3 12 1 US-08-439-817-18 Sequence 18, App1
139 27 27.3 12 1 US-08-439-817-26 Sequence 26, App1
140 27 27.3 12 1 US-08-383-474B-26 Sequence 26, App1
141 27 27.3 12 1 US-08-485-508-38 Sequence 38, App1
142 27 27.3 12 1 US-08-485-508-46 Sequence 46, App1
143 27 27.3 12 1 US-08-465-391A-21 Sequence 21, App1
144 27 27.3 12 2 US-08-464-538B-21 Sequence 21, App1
145 27 27.3 12 2 US-08-463-076E-65 Sequence 65, App1
146 27 27.3 12 4 US-09-823-823-44 Sequence 44, App1
147 27 27.3 12 4 US-09-428-082B-651 Sequence 651, App
148 27 27.3 15 5 PCT-US93-11703-40 Sequence 40, App1
149 27 27.3 15 6 5378464-14 Patent No. 5378464
150 27 27.3 16 2 US-08-476-062A-21 Sequence 21, App1

ALIGNMENTS

RESULT 1
US-09-441-502B-68
; Sequence 68, Application US/09441502B
; Patent No. 6455041
; GENERAL INFORMATION:
; APPLICANT: Dundar, Bonita S.
; TITLE OF INVENTION: IMMUNOGENIC EPITOPES OF THE HUMAN ZONA PELLUCIDA PROTEIN
; FILE REFERENCE: 12231.2USU1
; CURRENT APPLICATION NUMBER: US/09/441,502B
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-441-502B-68

Query Match 34.3%; Score 34; DB 4; Length 19;
Best Local Similarity 53.8%; Pred. No. 52;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Qy 1 SSGGLNNNOTQ 13
Db 1 SSGGIDNNYQ 13

RESULT 2
US-08-817-177-2
; Sequence 2, Application US/08817177
; Patent No. 6096314
; GENERAL INFORMATION:
; APPLICANT: COHEN, Irvin R.
; APPLICANT: ELIAS, Dana
; TITLE OF INVENTION: PEPTIDES AND PHARMACEUTICAL COMPOSITIONS
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Browdy and Neimark, P.L.L.C.
; STREET: 419 Seventh Street, N. W.
; CITY: Washington
; COUNTRY: US
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/817,177
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/12686
; FILING DATE: 10-OCTOBER-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: ISRAEL APP. NO. 111,196
; FILING DATE: 07-OCTOBER-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: COHEN=27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-817-177-2
Query Match 32.3%; Score 32; DB 3; Length 12;
Best Local Similarity 45.5%; Pred. No. 65;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Qy 2 SDGLNNNOTQ 12
Db 1 ASSLWNTQDQ 11
RESULT 3
PCT-US95-12686-2
; Sequence 2, Application PC/TUS9512686
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PEPTIDES AND PHARMACEUTICAL COMPOSITIONS
; TITLE OF INVENTION: COMPRISSING THEM
; NUMBER OF SEQUENCES: 16

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/12666
FILING DATE:
CLASSIFICATION:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-12686-2

Query March 32.3%; Score 32; DB 5; Length 12;
Best Local Similarity 45.5%; Pred. No. 65;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 SDGLMNNQTD 12
Db 1 ASLMTNQDTQ 11

RESULT 4
US-08-218-025A-166
Sequence 166, Application US/08218025A
Patent No. 5556744
GENERAL INFORMATION:
APPLICANT: Weiner, David B.
APPLICANT: Ugen, Kenneth E.
APPLICANT: Williams, William V.
TITLE OF INVENTION: Methods and Compositions for Diagnosing
TITLE OF INVENTION: and Treating Certain HIV Infected Patients
NUMBER OF SEQUENCES: 197
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: P.O. Box 457, 321 No. 5556744rlstowm Road
CITY: Spring House
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/218,025A
FILING DATE: 24-MAR-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/891,451
FILING DATE: 29-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: WST33A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 540-9206
TELEFAX: (215) 540-5818
INFORMATION FOR SEQ ID NO: 166:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-218-025A-166

Query March 32.3%; Score 32; DB 1; Length 19;

Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSDGLMNN 8
Db 1 SLDDIMNN 8

RESULT 5
US-08-277-660A-22
Sequence 22, Application US/08277660A
Patent No. 5702908
GENERAL INFORMATION:
APPLICANT: Picklesley, Steven M.
APPLICANT: Lane, David P.
TITLE OF INVENTION: Interruption of Binding of MDM2 and p53
TITLE OF INVENTION: Protein and Therapeutic Application Thereof
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/277,660A
FILING DATE: 20-JUL-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Walter H.
REGISTRATION NUMBER: 24,190
REFERENCE/DOCKET NUMBER: A-60244/WHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-277-660A-22

Query March 31.3%; Score 31; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.8e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SDGLM 6
Db 3 SDGLM 7

RESULT 6
US-08-424-957-7
Sequence 7, Application US/08424957
Patent No. 5770377
GENERAL INFORMATION:
APPLICANT: Picklesley, Steven M.
APPLICANT: Lane, David P.
TITLE OF INVENTION: Interruption of Binding of MDM2 and p53
TITLE OF INVENTION: Protein and Therapeutic Application Thereof
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco

STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/424,957
FILING DATE: 19-APR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/277,660
FILING DATE: 20-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Walter H.
REGISTRATION NUMBER: 24,190
REFERENCE/DOCKET NUMBER: A-61228/WHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
US-08-424-957-7

Query Match 31.3%; Score 31; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.8e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 SDGLW 6
Db 3 SDGLW 7

RESULT 7
US-09-035-686-7
Sequence 7, Application US/09035686
Patent No. 6153391
GENERAL INFORMATION:
APPLICANT: Pickelley, Steven M.
APPLICANT: Lane, David P.
TITLE OF INVENTION: Interruption of Binding of MDN2 and P53
TITLE OF INVENTION: Protein and Therapeutic Application Thereof
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Flehr, Hohbach, Teest, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/035,686
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/424,957
FILING DATE: 19-APR-1995
APPLICATION NUMBER: US 08/277,660
FILING DATE: 20-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Walter H.

REGISTRATION NUMBER: 24,190
REFERENCE/DOCKET NUMBER: A-61228/WHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
US-09-035-686-7

Query Match 31.3%; Score 31; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.8e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 SDGLW 6
Db 3 SDGLW 7

RESULT 8
US-08-592-500-6
Sequence 6, Application US/08592500
Patent No. 6005089
GENERAL INFORMATION:
APPLICANT: Lanza, Francois
APPLICANT: Phillips, David R.
APPLICANT: Cazenave, Jean-Pierre
TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend Kourile and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,500
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/089,455
FILING DATE: 09-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 12418-28
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..13
OTHER INFORMATION: /note= "K5/6 peptide residues
US-08-592-500-6
Query Match 30.3%; Score 30; DB 3; Length 13;

Best Local Similarity 71.4%; Pred. No. 1.5e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 QLFLEHS 18
Db 7 QLFLEHN 13

RESULT 9

US-08-195-006-6
Sequence 6, Application US/08195006
Patent No. 6083688
GENERAL INFORMATION:
APPLICANT: Lanza, Francois
APPLICANT: Phillips, David R.
TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/195,006
FILING DATE: 10-FEB-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/089,455
FILING DATE: 09-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 12418-28
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..13
OTHER INFORMATION: /note= "K5/6 peptide residues
US-08-195-006-6

Query Match 30.3%; Score 30; DB 3; Length 13;
Best Local Similarity 71.4%; Pred. No. 1.5e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 QLFLEHS 18
Db 7 QLFLEHN 13

RESULT 10
PCT-US94-07644A-6
Sequence 6, Application PC/TUS9407644A
GENERAL INFORMATION:
APPLICANT: COR Therapeutics, Inc.
TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses
NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/07644A
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 012418-003000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..13
OTHER INFORMATION: /note= "K5/6 peptide residues
PCT-US94-07644A-6

Query Match 30.3%; Score 30; DB 5; Length 13;
Best Local Similarity 71.4%; Pred. No. 1.5e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 QLFLEHS 18
Db 7 QLFLEHN 13

RESULT 11
US-09-638-203-25
Sequence 25, Application US/09638203
Patent No. 6602501
GENERAL INFORMATION:
APPLICANT: Daniel E.H. Afar
APPLICANT: Rene S. Hubert
APPLICANT: Aya Jakobovits
APPLICANT: Arthur B. Raitano
TITLE OF INVENTION: NOVEL C-TYPE LECTIN TRANSMEMBRANE
TITLE OF INVENTION: ANTIGEN EXPRESSED IN HUMAN PROSTATE CANCER AND USES THEREOF
FILE REFERENCE: 129 2015U1
CURRENT APPLICATION NUMBER: US/09/638,203
CURRENT FILING DATE: 2000-08-11
PRIOR APPLICATION NUMBER: 60/148,935
PRIOR FILING DATE: 1999-08-12
NUMBER OF SEQ ID NOS: 47
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 25
LENGTH: 14
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-638-203-25

Query Match 30.3%; Score 30; DB 4; Length 14;
Best Local Similarity 55.6%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 GLMNNOTQ 12

Db 1 GLWRNGDQ 9

RESULT 12

US-08-218-025A-82
; Sequence 82, Application US/08218025A
; Patent No. 5556744
; GENERAL INFORMATION:
; APPLICANT: Weiner, David B.
; APPLICANT: Ugen, Kenneth E.
; APPLICANT: Williams, William V.
; TITLE OF INVENTION: Methods and Compositions for Diagnosing
; TITLE OF INVENTION: and Treating Certain HIV Infected Patients
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: P.O. Box 457, 321 No. 5556744ristown Road
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/218.025A
; FILING DATE: 24-MAR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/891,451
; FILING DATE: 29-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST33A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 540-9206
; TELEFAX: (215) 540-5818
; INFORMATION FOR SEQ ID NO: 82:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-218-025A-82

Query Match 30.3%; Score 30; DB 1; Length 20;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 DGLWNN 8
| : |||
Db 1 DDWNN 6

RESULT 13

US-08-612-973-73
; Sequence 73, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHAYE P.C.
; STREET: 1100 NORTH GLEBE ROAD

CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612.973
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 73:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-612-973-73

Query Match 30.3%; Score 30; DB 3; Length 20;
Best Local Similarity 38.5%; Pred. No. 2.4e+02;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSDGLWNNOTOL 13
| : |||
Db 7 NINGQWHINSTAL 19

RESULT 14

US-08-612-973-82
; Sequence 82, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHAYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612.973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100

INFORMATION FOR SEQ ID NO: 82:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-612-973-82

Query Match 30.3%; Score 30; DB 3; Length 20;
Best Local Similarity 40.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 6 WNNQOTQFL 15
DB 7 WGENDTDFV 16

RESULT 15

US-08-927-597-73
Sequence 73, Application US/08927597
Patent No. 6245503
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BUYSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4100
TELEFAX: (703) 816-4100
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,597
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 73:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-927-597-73

Query Match 30.3%; Score 30; DB 3; Length 20;
Best Local Similarity 38.5%; Pred. No. 2.4e+02;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 1 SSDGLMNNQOTQL 13
DB 7 NTNGCMWHINSTAL 19

RESULT 16

US-08-927-597-82
Sequence 82, Application US/08927597
Patent No. 6245503
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BUYSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4100
TELEFAX: (703) 816-4100
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,597
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 82:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-927-597-82

Query Match 30.3%; Score 30; DB 3; Length 20;
Best Local Similarity 40.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 6 WNNQOTQFL 15
DB 7 WGENDTDFV 16

RESULT 17

US-08-635-886C-38
Sequence 38, Application US/08635886C
Patent No. 6555114
GENERAL INFORMATION:
APPLICANT: LEROUX-ROELS, Geert
APPLICANT: DELEYS, Robert
APPLICANT: MAERTENS, Geert
TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
FILE REFERENCE: 2752-18
CURRENT APPLICATION NUMBER: US/08/635,886C
PRIOR APPLICATION NUMBER: PCT/EP94/03555
PRIOR FILING DATE: 1994-10-28
PRIOR APPLICATION NUMBER: EP 93402718.6

```
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38
; LENGTH: 20
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-635-886C-38
```

```
Query Match      30.3%; Score 30; DB 4; Length 20;
Best Local Similarity 40.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
QY      6 WNNNQIOLFL 15
        | | | | |
Db      7 WGENDTDFV 16
```

```
RESULT 18
US-08-635-886C-47
; Sequence 47, Application US/08635886C
; Patent No. 6555114
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; FILE REFERENCE: 2752-18
; CURRENT APPLICATION NUMBER: US/08/635,886C
; CURRENT FILING DATE: 1996-04-25
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47
; LENGTH: 20
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-635-886C-47
```

```
Query Match      30.3%; Score 30; DB 4; Length 20;
Best Local Similarity 38.5%; Pred. No. 2.4e+02;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
```

```
QY      1 SSDGLMNNQIOL 13
        : | | | | |
Db      7 NTNGSMHINSTAL 19
```

```
RESULT 19
US-08-635-886C-56
; Sequence 56, Application US/08635886C
; Patent No. 6555114
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MARTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; FILE REFERENCE: 2752-18
; CURRENT APPLICATION NUMBER: US/08/635,886C
; CURRENT FILING DATE: 1996-04-25
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56
; LENGTH: 20
```

```
; TYPE: PRT
; ORGANISM: hepatitis C virus
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (2)..(2)
; OTHER INFORMATION: Xaa is Asn, Lys, Asp or Arg
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (3)..(3)
; OTHER INFORMATION: Xaa is Val, Ile or Leu
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (4)..(4)
; OTHER INFORMATION: Xaa is Gln, Ser or Tyr
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (6)..(6)
; OTHER INFORMATION: Xaa is Ile or Val
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (8)..(8)
; OTHER INFORMATION: Xaa is Thr or Ser
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (14)..(14)
; OTHER INFORMATION: Xaa is Leu or Ile
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (16)..(16)
; OTHER INFORMATION: Xaa is Ser or Arg
US-08-635-886C-56
```

```
Query Match      30.3%; Score 30; DB 4; Length 20;
Best Local Similarity 38.5%; Pred. No. 2.4e+02;
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 SSDGLMNNQIOL 13
        : | | | | |
Db      7 NXNGSMHKNXTAL 19
```

```
RESULT 20
US-08-974-690C-38
; Sequence 38, Application US/08974690C
; Patent No. 6613333
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/08/974,690C
; CURRENT FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38
; LENGTH: 20
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-974-690C-38
```

```
Query Match      30.3%; Score 30; DB 4; Length 20;
Best Local Similarity 40.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
QY      6 WNNNQIOLFL 15
        | | | | |
Db      7 WGENDTDFV 16
```

RESULT 21
US-08-974-690C-47
Sequence 47, Application US/08974690C
Patent No. 661333
GENERAL INFORMATION:
APPLICANT: LEROUX-ROELS, Geert
APPLICANT: DELEYS, Robert
APPLICANT: MAERTENS, Geert
TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
FILE REFERENCE: 2551-94
CURRENT APPLICATION NUMBER: US/08/974,690C
CURRENT FILING DATE: 1997-11-19
PRIOR APPLICATION NUMBER: PCT/EP94/03555
PRIOR FILING DATE: 1994-10-28
PRIOR APPLICATION NUMBER: EP 93402718.6
PRIOR FILING DATE: 1993-11-04
NUMBER OF SEQ ID NOS: 286
SOFTWARE: Patentin version 3.1
SEQ ID NO 47
LENGTH: 20
TYPE: PRT
ORGANISM: hepatitis C virus
US-08-974-690C-47

Query Match 30.3%; Score 30; DB 4; Length 20;
Best Local Similarity 38.5%; Pred. No. 2.4e+02;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSDGLMNNNOTOL 13
DB 7 NINGSMWHINSTAL 19

RESULT 22
US-08-974-690C-56
Sequence 56, Application US/08974690C
Patent No. 661333
GENERAL INFORMATION:
APPLICANT: LEROUX-ROELS, Geert
APPLICANT: DELEYS, Robert
APPLICANT: MAERTENS, Geert
TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
FILE REFERENCE: 2551-94
CURRENT APPLICATION NUMBER: US/08/974,690C
CURRENT FILING DATE: 1997-11-19
PRIOR APPLICATION NUMBER: PCT/EP94/03555
PRIOR FILING DATE: 1994-10-28
PRIOR APPLICATION NUMBER: EP 93402718.6
PRIOR FILING DATE: 1993-11-04
NUMBER OF SEQ ID NOS: 286
SOFTWARE: Patentin version 3.1
SEQ ID NO 56
LENGTH: 20
TYPE: PRT
ORGANISM: hepatitis C virus
FEATURE: MISC FEATURE
NAME/KEY: MISC FEATURE
LOCATION: (2)..(2)
OTHER INFORMATION: Xaa is Asn, Lys, Asp or Arg
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (3)..(3)
OTHER INFORMATION: Xaa is Val, Ile or Leu
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (4)..(4)
OTHER INFORMATION: Xaa is Gln, Ser or Tyr
FEATURE:
NAME/KEY: MISC FEATURE

LOCATION: (6)..(6)
OTHER INFORMATION: Xaa is Ile or Val
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (8)..(8)
OTHER INFORMATION: Xaa is Thr or Ser
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (14)..(14)
OTHER INFORMATION: Xaa is Leu or Ile
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (16)..(16)
OTHER INFORMATION: Xaa is Ser or Arg
US-08-974-690C-56

Query Match 30.3%; Score 30; DB 4; Length 20;
Best Local Similarity 38.5%; Pred. No. 2.4e+02;
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 SSDGLMNNNOTOL 13
DB 7 NXNGSMWHXNTAL 19

RESULT 23
US-09-881-710-15
Sequence 15, Application US/09881710
Patent No. 6673895
GENERAL INFORMATION:
APPLICANT: DESPRES, Philippe
APPLICANT: COURAGEOT, Marie-Pierre
APPLICANT: DEUBEL, Vincent
APPLICANT: CATTEAU, Adeline
TITLE OF INVENTION: PRO-APOPTOTIC FRAGMENTS OF THE DENGUE VIRUS ENVELOPE GLYCOPROTEIN:
FILE REFERENCE: 209671US0
CURRENT APPLICATION NUMBER: US/09/881,710
CURRENT FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 60/212,129
PRIOR FILING DATE: 2000-06-16
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Patentin version 3.1
SEQ ID NO 15
LENGTH: 20
TYPE: PRT
ORGANISM: Dengue virus
US-09-881-710-15

Query Match 30.3%; Score 30; DB 4; Length 20;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSDGLMNNNO 10
DB 10 SSEGAMKHAQ 19

RESULT 24
US-09-881-710-17
Sequence 17, Application US/09881710
Patent No. 6673895
GENERAL INFORMATION:
APPLICANT: DESPRES, Philippe
APPLICANT: COURAGEOT, Marie-Pierre
APPLICANT: DEUBEL, Vincent
APPLICANT: CATTEAU, Adeline
TITLE OF INVENTION: PRO-APOPTOTIC FRAGMENTS OF THE DENGUE VIRUS ENVELOPE GLYCOPROTEIN:
FILE REFERENCE: 209671US0
CURRENT APPLICATION NUMBER: US/09/881,710
CURRENT FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 60/212,129
PRIOR FILING DATE: 2000-06-16
NUMBER OF SEQ ID NOS: 30

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Dengue virus
US-09-881-710-17

Query Match 30.3%; Score 30; DB 4; Length 20;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSGGLMNNQ 10
Db 10 SSEGAMKHAQ 19

RESULT 25
US-08-974-685-47
; Sequence 47, Application US/08974685
; Patent No. 6689368
; GENERAL INFORMATION:

APPLICANT: LEROUX-ROELS, GEERT
DELEYS, ROBERT

TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF
HEPATITIS C VIRUS STRUCTURAL PROTEINS, PROPHYLACTIC AND
THERAPEUTIC COMPOSITIONS CONTAINING THE SAME

NUMBER OF SEQUENCES: 181

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHYE P.C.

STREET: 1100 NORTH GLEBE ROAD

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: U.S.A.

ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/974,685

FILING DATE: 19-No. 6689368-1997

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: BYRNE, THOMAS E.

REGISTRATION NUMBER: 32,205

REFERENCE/DOCKET NUMBER: 1487-11

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 816-4000

TELEFAX: (703) 816-4100

INFORMATION FOR SEQ ID NO: 47:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULAR TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 47:

US-08-974-685-47

Query Match 30.3%; Score 30; DB 4; Length 20;
Best Local Similarity 38.5%; Pred. No. 2.4e+02;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSGGLMNNQ 13
Db 7 NTNGSMHNSYAL 19

RESULT 26
US-08-974-685-56
; Sequence 56, Application US/08974685

; Patent No. 6689368
; GENERAL INFORMATION:
APPLICANT: LEROUX-ROELS, GEERT
DELEYS, ROBERT
MAERTENS, GEERT

TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF
HEPATITIS C VIRUS STRUCTURAL PROTEINS, PROPHYLACTIC AND
THERAPEUTIC COMPOSITIONS CONTAINING THE SAME

NUMBER OF SEQUENCES: 181

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHYE P.C.

STREET: 1100 NORTH GLEBE ROAD

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: U.S.A.

ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/974,685

FILING DATE: 19-No. 6689368-1997

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: BYRNE, THOMAS E.

REGISTRATION NUMBER: 32,205

REFERENCE/DOCKET NUMBER: 1487-11

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 816-4000

TELEFAX: (703) 816-4100

INFORMATION FOR SEQ ID NO: 56:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULAR TYPE: peptide

FEATURE:

NAME/KEY: misc-feature

LOCATION: 2

OTHER INFORMATION: Xaa is Asn or Lys or Asp or Arg

FEATURE:

NAME/KEY: misc-feature

LOCATION: 3

OTHER INFORMATION: Xaa is Val or Ile or Leu

FEATURE:

NAME/KEY: misc-feature

LOCATION: 4

OTHER INFORMATION: Xaa is Gln or Ser or Tyr

FEATURE:

NAME/KEY: misc-feature

LOCATION: 6

OTHER INFORMATION: Xaa is Ile or Val

FEATURE:

NAME/KEY: misc-feature

LOCATION: 8

OTHER INFORMATION: Xaa is Thr or Ser

FEATURE:

NAME/KEY: misc-feature

LOCATION: 14

OTHER INFORMATION: Xaa is Leu or Ile

FEATURE:

NAME/KEY: misc-feature

LOCATION: 16

OTHER INFORMATION: Xaa is Ser or Arg

SEQUENCE DESCRIPTION: SEQ ID NO: 56:

US-08-974-685-56

Query Match 30.3%; Score 30; DB 4; Length 20;
Best Local Similarity 38.5%; Pred. No. 2.4e+02;
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 SSDGLMNNOTOL 13
DB 7 NXNGSMHXKXTAL 19

RESULT 27

US-08-014-979-117
Sequence 117, Application US/08014979

Patent No. 5510240

GENERAL INFORMATION:

APPLICANT: Lam, Kit S. et al.

TITLE OF INVENTION: Random Bio-Oligomer Library, A Method of

TITLE OF INVENTION: Synthesis Thereof, and a Method of Use Thereof

NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/014,979

FILING DATE: 19930208

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Mierock, S. Leslie

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 7156-041

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212 790-9090

TELEFAX: 212 869-8864/9741

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 117:

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-08-014-979-117

Query Match

Best Local Similarity 100.0%; Pred. No. 3.8e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 WNNN 9

DB 1 WNNN 4

RESULT 28

US-08-292-968-43

Sequence 43, Application US/08292968

Patent No. 5856122

GENERAL INFORMATION:

APPLICANT: READ, Randy J.

APPLICANT: STEIN, Penelope E.

APPLICANT: COCKLE, Stephen A.

APPLICANT: COOMEN, Raymond P.

APPLICANT: LOOSMORE, Sheena

APPLICANT: KLEIN, Michel H.

APPLICANT: ARMSTRONG, Glen D.

APPLICANT: HAZES, Bart

TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Sim & McBurney

STREET: Suite 701, 330 University Avenue

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: MSG 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/292,968

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/110,947

FILING DATE: 24-AUG-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/251,121

FILING DATE: 31-MAY-1994

ATTORNEY/AGENT INFORMATION:

NAME: STEWART, Michael I.

REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-388

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 43:

SEQUENCE CHARACTERISTICS:

LENGTH: 7 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-292-968-43

Query Match

Best Local Similarity 66.7%; Pred. No. 3.8e+05;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 DGLMNN 8

DB 2 NGLMND 7

RESULT 29

US-08-467-974-43

Sequence 43, Application US/08467974

Patent No. 5965385

GENERAL INFORMATION:

APPLICANT: READ, Randy J.

APPLICANT: STEIN, Penelope E.

APPLICANT: COCKLE, Stephen A.

APPLICANT: COOMEN, Raymond P.

APPLICANT: LOOSMORE, Sheena

APPLICANT: KLEIN, Michel H.

APPLICANT: ARMSTRONG, Glen D.

APPLICANT: HAZES, Bart

TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Sim & McBurney

STREET: Suite 701, 330 University Avenue

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: MSG 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/467,974
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/467,536
FILING DATE: 06-JUN-1995
APPLICATION NUMBER: US 08/292,968
FILING DATE: 22-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/251,121
FILING DATE: 31-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/110,947
FILING DATE: 24-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-454 MIS:V9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-467-974-43

Query Match 29.3%; Score 29; DB 2; Length 7;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 DGLMNN 8
:||||:
Db 2 NGLMND 7

RESULT 30
US-08-467-536-43
Sequence 43, Application US/08467536
Patent No. 5977304
GENERAL INFORMATION:
APPLICANT: READ, Randy J.
APPLICANT: STEIN, Penelope E.
APPLICANT: COCKLE, Stephen A.
APPLICANT: OOMEN, Raymond P.
APPLICANT: LOOSMORE, Sheena
APPLICANT: KLEIN, Michel H.
APPLICANT: ARMSTRONG, Glen D.
TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,536
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/292,968
FILING DATE: 22-AUG-1994
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/251,121
FILING DATE: 31-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/110,947
FILING DATE: 24-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-455 MIS:V9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-467-536-43

Query Match 29.3%; Score 29; DB 2; Length 7;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 DGLMNN 8
:||||:
Db 2 NGLMND 7

RESULT 31
US-08-467-976-43
Sequence 43, Application US/08467976
Patent No. 6018022
GENERAL INFORMATION:
APPLICANT: READ, Randy J.
APPLICANT: STEIN, Penelope E.
APPLICANT: COCKLE, Stephen A.
APPLICANT: OOMEN, Raymond P.
APPLICANT: LOOSMORE, Sheena
APPLICANT: KLEIN, Michel H.
APPLICANT: ARMSTRONG, Glen D.
TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,976
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/292,968
FILING DATE: 22-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/251,121
FILING DATE: 31-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/110,947
FILING DATE: 24-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-453 MIS:V9

TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-467-976-43

Query Match 29.3% Score 29; DB 3; Length 7;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 3 DGLMNN 8
Db 2 NGLMND 7

RESULT 32
US-09-082-514-43
Sequence 43, Application US/09082514
Patent No. 6168928
GENERAL INFORMATION:

APPLICANT: READ, Randy J.
APPLICANT: STEIN, Penelope E.
APPLICANT: COCKLE, Stephen A.
APPLICANT: OOMEN, Raymond P.
APPLICANT: KLEIN, Michel H.
APPLICANT: ARMSTRONG, Glen D.
APPLICANT: HAZES, Bart

TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/082,514
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/292,968
FILING DATE: 24-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-810
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-082-514-43

Query Match 29.3% Score 29; DB 3; Length 7;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 3 DGLMNN 8

Db 2 NGLMND 7

RESULT 33
US-09-563-222C-53
Sequence 53, Application US/09563222C
Patent No. 6696620
GENERAL INFORMATION:

APPLICANT: EPICYTE PHARMACEUTICALS, INC.
APPLICANT: HIATT, ANDREW C.
APPLICANT: HEIN, MICHAEL B.

TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN ARRAYS IN PLANT CELLS
FILE REFERENCE: 068904-0501
CURRENT APPLICATION NUMBER: US/09/563,222C
CURRENT FILING DATE: 2000-05-02
PRIOR APPLICATION NUMBER: PCT/US01/14349
PRIOR FILING DATE: 2001-05-02
PRIOR APPLICATION NUMBER: 09/563,222
PRIOR FILING DATE: 2000-05-02
NUMBER OF SEQ ID NOS: 182
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 53
LENGTH: 7
TYPE: PRT
ORGANISM: Mus musculus
US-09-563-222C-53

Query Match 29.3% Score 29; DB 4; Length 7;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 4 GLMNN 9
Db 2 GYMNS 7

RESULT 34
US-09-570-921-141
Sequence 141, Application US/09570921
Patent No. 6455265
GENERAL INFORMATION:

APPLICANT: SERRES, PIERRE-FRANCOIS
TITLE OF INVENTION: METHOD FOR OBTAINING VACCINES FOR PREVENTING THE
TITLE OF INVENTION: PATHOGENIC EFFECTS RELATED TO A RETROVIRAL INFECTION
FILE REFERENCE: 106213
CURRENT APPLICATION NUMBER: US/09/570,921

CURRENT FILING DATE: 2000-05-15
PRIOR APPLICATION NUMBER: PCT/FR98/02447
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: FR/97/14387
PRIOR FILING DATE: 1997-11-17
NUMBER OF SEQ ID NOS: 144
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 141
LENGTH: 14
TYPE: PRT
ORGANISM: Human
US-09-570-921-141

Query Match 29.3% Score 29; DB 4; Length 14;
Best Local Similarity 75.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 10 QTOLEH 17
Db 1 KTOLEH 8

RESULT 35
US-10-032-330-52
Sequence 52, Application US/10032330
Patent No. 6770626

GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel
TITLE OF INVENTION: Tissue Remodeling
FILE REFERENCE: BEN-SASSON=7
CURRENT APPLICATION NUMBER: US/10/032,330
CURRENT FILING DATE: 2001-12-31
PRIOR APPLICATION NUMBER: PCT/US00/32852
PRIOR FILING DATE: 2000-12-04
PRIOR APPLICATION NUMBER: US 09/161,094
PRIOR FILING DATE: 1998-09-25
NUMBER OF SEQ ID NOS: 59
SOFTWARE: Patent version 3.1
SEQ ID NO: 52
LENGTH: 17
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic
US-10-032-330-52

Query Match 29.3%; Score 29; DB 4; Length 17;
Best Local Similarity 50.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 SDGLMNNQTOLEFL 15
Db 1 SDMTSRNSSTQLWL 14

RESULT 36
US-08-190-788A-19
Sequence 19, Application US/08190788A
Patent No. 5608035
GENERAL INFORMATION:
APPLICANT: Yanofsky, Stephen D.
APPLICANT: Barrett, Ronald W.
APPLICANT: Baldwin, David N.
APPLICANT: Jacobs, Jeff W.
TITLE OF INVENTION: Peptides and Compounds That Bind to the
TITLE OF INVENTION: IL-1 Receptor
NUMBER OF SEQUENCES: 312
CORRESPONDENCE ADDRESS:
ADDRESSEE: Affymax Technologies N.V.
STREET: 4001 Miranda Avenue
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/190,788A
FILING DATE: 02-FEB-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847,567
FILING DATE: 05-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Stevens, Lauren L.
REGISTRATION NUMBER: 36,691
REFERENCE/DOCKET NUMBER: 1019.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-496-2300
TELEFAX: 415-424-0832
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
US-08-190-788A-19

Query Match 28.3%; Score 28; DB 1; Length 8;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSDGLW 6
Db 3 TDDGLW 8

RESULT 37
US-08-383-474B-24
Sequence 24, Application US/08383474B
Patent No. 5767234
GENERAL INFORMATION:
APPLICANT: Yanofsky, Stephen D.
APPLICANT: Barrett, Ronald W.
APPLICANT: Baldwin, David N.
APPLICANT: Jacobs, Jeff W.
TITLE OF INVENTION: Peptides and Compounds That Bind to
TITLE OF INVENTION: the IL-1 Receptor
NUMBER OF SEQUENCES: 314
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend & Townsend & Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/383,474B
FILING DATE: 01-FEB-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/190,788
FILING DATE: 02-FEB-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Stevens, Lauren L.
REGISTRATION NUMBER: 36,691
REFERENCE/DOCKET NUMBER: 1019.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-496-2300
TELEFAX: 415-424-0832
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-383-474B-24

Query Match 28.3%; Score 28; DB 1; Length 8;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSDGLW 6
Db 3 TDDGLW 8

RESULT 38
US-08-465-391A-19
Sequence 19, Application US/08465391A
Patent No. 5786531

GENERAL INFORMATION:
APPLICANT: Barrett, Ronald W.
APPLICANT: Yanofsky, Stephen D.
APPLICANT: Baldwin, David
APPLICANT: Jacobs, Jeff W.
APPLICANT: Bovy, Philippe R.
APPLICANT: Leahy, Ellen M.
APPLICANT: Pottorf, Richard S.
TITLE OF INVENTION: Peptides and Compounds That Bind to the
NUMBER OF INVENTION: IL-1 Receptor
NUMBER OF SEQUENCES: 405
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew
STREET: One Market Plaza, Steuart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,391A
FILING DATE: 05-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/373,474
FILING DATE: 01-FEB-1995
CLASSIFICATION: 514
REGISTRATION NUMBER: 32,483
REFERENCE/DOCKET NUMBER: 16528A-001840/1019.2A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2422
FAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-465-391A-19

Query Match 28.3%; Score 28; DB 1; Length 8;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSDGLW 6
: ||||
Db 3 TDDGLW 8

RESULT 39
US-08-464-538B-19
Sequence 19, Application US/08464538B
Patent No. 5861476
GENERAL INFORMATION:
APPLICANT: Barrett, Ronald W.
APPLICANT: Yanofsky, Stephen D.
APPLICANT: Baldwin, David
APPLICANT: Jacobs, Jeff W.
APPLICANT: Bovy, Philippe R.
APPLICANT: Leahy, Ellen M.
APPLICANT: Pottorf, Richard S.
TITLE OF INVENTION: Peptides and Compounds That Bind to the

TITLE OF INVENTION: IL-1 Receptor
NUMBER OF SEQUENCES: 402
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/464,538B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/373,474
FILING DATE: 01-FEB-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/190,788
FILING DATE: 02-FEB-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 16528A-001810
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
FAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-464-538B-19

Query Match 28.3%; Score 28; DB 2; Length 8;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSDGLW 6
: ||||
Db 3 TDDGLW 8

RESULT 40
US-08-463-076E-63
Sequence 63, Application US/08463076E
Patent No. 5880096
GENERAL INFORMATION:
APPLICANT: Barrett, Ronald W.
APPLICANT: Yanofsky, Stephen D.
TITLE OF INVENTION: Peptides and Compounds That Bind to the
NUMBER OF SEQUENCES: 392
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,076E
FILING DATE: 05-JUN-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Snyder, Joseph R.
REGISTRATION NUMBER: 39,381
REFERENCE/DOCKET NUMBER: 16528A-001850US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-463-076E-63

```

```

Query Match      28.3% Score 28; DB 2; Length 8;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      1 SSDGLW 6
        : ||||
Db      3 TDDGLW 8

```

Search completed: December 30, 2004, 16:32:49
 Job time : 39 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 16:31:35 ; Search time 141 Seconds
(without alignments)
45.923 Million cell updates/sec

Title: US-10-718-321-1

Perfect score: 99

Sequence: 1 SSDGLMNNNQTLFLEHS 18

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1599051 seqs, 359727711 residues

Total number of hits satisfying chosen parameters: 292596

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 100%

Listing first 150 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09C_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	34.3	12	US-10-649-873-131	Sequence 131, App
2	34.3	20	US-10-225-567A-1690	Sequence 1690, App
3	33.3	18	US-10-418-820-18	Sequence 18, Appl
4	33.3	20	US-09-974-879-288	Sequence 288, App
5	33.3	20	US-09-305-736-276	Sequence 276, App
6	33.3	20	US-09-818-683-276	Sequence 276, App
7	33.3	20	US-09-818-683-276	Sequence 276, App
8	33.3	20	US-10-621-401-288	Sequence 288, App
9	31.3	19	US-10-102-143-28	Sequence 28, Appl
10	31.3	20	US-10-029-386-31498	Sequence 31498, A
11	30.3	12	US-09-805-290A-15	Sequence 15, Appl
12	30.3	13	US-10-212-499-6	Sequence 6, Appl
13	30.3	14	US-09-932-613-254	Sequence 254, App

14	30	30.3	14	US-09-932-322-254	Sequence 254, App
15	30	30.3	14	US-10-460-512-25	Sequence 25, Appl
16	30	30.3	17	US-10-466-085A-16	Sequence 16, Appl
17	30	30.3	20	US-09-881-710-15	Sequence 15, Appl
18	30	30.3	20	US-09-881-710-17	Sequence 17, Appl
19	30	30.3	20	US-09-973-025-73	Sequence 73, Appl
20	30	30.3	20	US-09-973-025-82	Sequence 82, Appl
21	30	30.3	20	US-09-899-103-73	Sequence 73, Appl
22	30	30.3	20	US-09-899-103-82	Sequence 82, Appl
23	30	30.3	20	US-09-995-808-82	Sequence 82, Appl
24	30	30.3	20	US-09-995-808-82	Sequence 82, Appl
25	30	30.3	20	US-09-995-860-73	Sequence 73, Appl
26	30	30.3	20	US-09-995-860-82	Sequence 82, Appl
27	30	30.3	20	US-09-995-791-73	Sequence 73, Appl
28	30	30.3	20	US-09-995-791-82	Sequence 82, Appl
29	30	30.3	20	US-10-651-165-38	Sequence 38, Appl
30	30	30.3	20	US-10-651-165-47	Sequence 47, Appl
31	30	30.3	20	US-10-651-165-56	Sequence 56, Appl
32	30	30.3	20	US-10-634-895-15	Sequence 15, Appl
33	30	30.3	20	US-10-634-895-17	Sequence 17, Appl
34	30	30.3	20	US-10-311-213-22	Sequence 22, Appl
35	30	30.3	20	US-10-321-798-73	Sequence 73, Appl
36	30	30.3	20	US-10-321-798-82	Sequence 82, Appl
37	29	29.3	7	US-09-563-222-53	Sequence 53, Appl
38	29	29.3	7	US-10-783-950-53	Sequence 53, Appl
39	29	29.3	15	US-10-416-249-557	Sequence 557, App
40	29	29.3	15	US-10-416-249-563	Sequence 563, App
41	29	29.3	15	US-10-416-249-564	Sequence 564, App
42	29	29.3	15	US-10-416-249-565	Sequence 565, App
43	29	29.3	16	US-09-995-529-111	Sequence 111, App
44	29	29.3	16	US-09-995-529-111	Sequence 111, App
45	29	29.3	16	US-10-676-809-4	Sequence 4, Appl
46	29	29.3	17	US-10-032-330-52	Sequence 52, Appl
47	29	29.3	17	US-10-258-144-472	Sequence 472, App
48	29	29.3	17	US-10-882-640-52	Sequence 52, Appl
49	28.5	28.8	20	US-10-776-013-647	Sequence 647, App
50	28	28.3	8	US-10-609-217-649	Sequence 649, App
51	28	28.3	8	US-10-632-188-649	Sequence 649, App
52	28	28.3	8	US-10-651-723-649	Sequence 649, App
53	28	28.3	8	US-10-645-723-649	Sequence 649, App
54	28	28.3	8	US-10-666-996-649	Sequence 649, App
55	28	28.3	8	US-10-653-048-649	Sequence 649, App
56	28	28.3	9	US-09-940-127B-42	Sequence 42, Appl
57	28	28.3	11	US-10-149-135-761	Sequence 761, Appl
58	28	28.3	11	US-10-149-135-1907	Sequence 1907, App
59	28	28.3	13	US-10-224-999A-390	Sequence 590, App
60	28	28.3	13	US-10-467-209-7	Sequence 7, Appl
61	28	28.3	13	US-10-468-496-52	Sequence 52, Appl
62	28	28.3	13	US-10-468-496-52	Sequence 52, Appl
63	28	28.3	13	US-10-468-496-52	Sequence 52, Appl
64	28	28.3	14	US-09-966-422B-60	Sequence 60, Appl
65	28	28.3	14	US-09-932-613-198	Sequence 198, App
66	28	28.3	14	US-09-932-613-224	Sequence 224, App
67	28	28.3	14	US-09-932-613-234	Sequence 234, App
68	28	28.3	14	US-09-932-322-198	Sequence 198, App
69	28	28.3	14	US-09-932-322-224	Sequence 224, App
70	28	28.3	14	US-09-932-322-234	Sequence 234, App
71	28	28.3	14	US-10-158-847-130	Sequence 130, App
72	28	28.3	14	US-10-262-772A-60	Sequence 60, Appl
73	28	28.3	14	US-10-224-999A-600	Sequence 600, Appl
74	28	28.3	14	US-10-224-999A-601	Sequence 601, App
75	28	28.3	14	US-10-158-825-130	Sequence 130, App
76	28	28.3	14	US-10-158-825-130	Sequence 130, App
77	28	28.3	14	US-10-665-478-757	Sequence 757, App
78	28	28.3	15	US-09-953-510-73	Sequence 73, Appl
79	28	28.3	15	US-09-953-510-73	Sequence 73, Appl
80	28	28.3	15	US-10-147-255-73	Sequence 73, Appl
81	28	28.3	15	US-10-224-999A-611	Sequence 611, App
82	28	28.3	15	US-10-224-999A-612	Sequence 612, App
83	28	28.3	15	US-10-224-999A-613	Sequence 613, App
84	28	28.3	15	US-10-149-135-1914	Sequence 1914, App
85	28	28.3	15	US-10-149-135-1970	Sequence 1970, App
86	28	28.3	15	US-10-149-135-2019	Sequence 2019, App

PRIOR APPLICATION NUMBER: 60/374,066
PRIOR FILING DATE: 2002-04-19
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn version 3.2
SEQ ID NO: 18
LENGTH: 18
TYPE: PRT
ORGANISM: Candida tropicalis
US-10-418-820-18

Query Match 33.3%; Score 33; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GLMNN 8
DB 1 GLMNN 5

RESULT 4

US-09-974-879-288
Sequence 288, Application US/09974879
Publication No. US20030028003A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 125 Human Secreted Proteins
FILE REFERENCE: P2020P2
CURRENT APPLICATION NUMBER: US/09/974,879
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: US 60/239,893
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: US 09/818,683
PRIOR FILING DATE: 2001-03-28
PRIOR APPLICATION NUMBER: US 09/305,736
PRIOR FILING DATE: 1999-05-05
PRIOR APPLICATION NUMBER: PCT/US98/23435
PRIOR FILING DATE: 1998-11-04
PRIOR APPLICATION NUMBER: US 60/064,911
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: US 60/064,912
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: US 60/064,983
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: US 60/064,900
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: US 60/064,988
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: US 60/064,987
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: US 60/064,908
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: US 60/064,984
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: US 60/064,985
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: US 60/066,094
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: US 60/066,100
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: US 60/066,089
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: US 60/066,095
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: US 60/066,090
PRIOR FILING DATE: 1997-11-17
NUMBER OF SEQ ID NOS: 611
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 288
LENGTH: 20
TYPE: PRT
ORGANISM: Homo sapiens
US-09-974-879-288

Query Match 33.3%; Score 33; DB 10; Length 20;
Best Local Similarity 50.0%; Pred. No. 3.3e+02;
Matches 17; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 5 LMMNNOTQFLFHS 18
DB 5 LVTNNOTQRLQEA 18

RESULT 5

US-09-305-736-276
Sequence 276, Application US/09305736
Publication No. US20030088078A1
GENERAL INFORMATION:
APPLICANT: Feng et al.
TITLE OF INVENTION: 125 Human Secreted Proteins
FILE REFERENCE: P2020P1
CURRENT APPLICATION NUMBER: US/09/305,736
CURRENT FILING DATE: 1999-05-05
PRIOR APPLICATION NUMBER: PCT/US98/23435
PRIOR FILING DATE: 1998-11-04
PRIOR APPLICATION NUMBER: 60/064,911
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/064,912
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/064,983
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/064,900
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/064,988
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/064,987
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/064,908
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/064,984
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/064,985
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/066,094
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: 60/066,100
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: 60/066,089
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: 60/066,095
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: 60/066,090
PRIOR FILING DATE: 1997-11-17
NUMBER OF SEQ ID NOS: 612
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 276
LENGTH: 20
TYPE: PRT
ORGANISM: Homo sapiens
US-09-305-736-276

Query Match 33.3%; Score 33; DB 10; Length 20;
Best Local Similarity 50.0%; Pred. No. 3.3e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 5 LMMNNOTQFLFHS 18
DB 5 LVTNNOTQRLQEA 18

RESULT 6
US-09-818-683-276
Sequence 276, Application US/09818683
Publication No. US20030211472A1
GENERAL INFORMATION:
APPLICANT: Feng et al.
TITLE OF INVENTION: 125 Human Secreted Proteins

FILE REFERENCE: P2020P1
CURRENT APPLICATION NUMBER: US/09/818,683
CURRENT FILING DATE: 2001-03-28
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 612
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 276
LENGTH: 20
TYPE: PRT
ORGANISM: Homo sapiens
US-09-818-683-276

Query Match 33.3%; Score 33; DB 10; Length 20;
Best Local Similarity 50.0%; Pred. No. 3.3e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 5 LMNNQOTQLEHNS 18
| | | | | : :
Db 5 LVTNNQOTQLQEA 18

RESULT 7
US-09-818-683-276
Sequence 276, Application US/09818683
Publication No. US20040185440A9
GENERAL INFORMATION:
APPLICANT: Feng et al.
TITLE OF INVENTION: 125 Human Secreted Proteins
FILE REFERENCE: P2020P1
CURRENT APPLICATION NUMBER: US/09/818,683
CURRENT FILING DATE: 2001-03-28
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 612
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 276
LENGTH: 20
TYPE: PRT
ORGANISM: Homo sapiens
US-09-818-683-276

Query Match 33.3%; Score 33; DB 11; Length 20;
Best Local Similarity 50.0%; Pred. No. 3.3e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 5 LMNNQOTQLEHNS 18
| | | | | : :
Db 5 LVTNNQOTQLQEA 18

RESULT 8
US-10-621-401-288
Sequence 288, Application US/10621401
Publication No. US20040038277A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 125 Human Secreted Proteins
FILE REFERENCE: P2020P2C1
CURRENT APPLICATION NUMBER: US/10/621,401
CURRENT FILING DATE: 2003-07-18
PRIOR APPLICATION NUMBER: US 09/974,879
PRIOR FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: US 60/239,893
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: US 09/818,683
PRIOR FILING DATE: 2001-03-28
PRIOR APPLICATION NUMBER: US 09/305,736
PRIOR FILING DATE: 1999-05-05
PRIOR APPLICATION NUMBER: PCT/US98/23435
PRIOR FILING DATE: 1998-11-04
PRIOR APPLICATION NUMBER: US 60/064,911
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: US 60/064,912
PRIOR FILING DATE: 1997-11-07

PRIOR APPLICATION NUMBER: US 60/064,983
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: US 60/064,900
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: US 60/064,988
PRIOR FILING DATE: 1997-11-07
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 611
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 288
LENGTH: 20
TYPE: PRT
ORGANISM: Homo sapiens
US-10-621-401-288

Query Match 33.3%; Score 33; DB 15; Length 20;
Best Local Similarity 50.0%; Pred. No. 3.3e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 5 LMNNQOTQLEHNS 18
| | | | | : :
Db 5 LVTNNQOTQLQEA 18

RESULT 9
US-10-102-143-28
Sequence 28, Application US/10102143
Publication No. US20030185851A1
GENERAL INFORMATION:
APPLICANT: Soldati, Dominique
APPLICANT: Weisner, Markus
TITLE OF INVENTION: TET transactivator system
FILE REFERENCE: 04630/016001
CURRENT APPLICATION NUMBER: US/10/102,143
CURRENT FILING DATE: 2003-01-21
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 28
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: TATi-3
OTHER INFORMATION: presumed amino acid sequence of activating domain
US-10-102-143-28

Query Match 31.3%; Score 31; DB 14; Length 19;
Best Local Similarity 60.0%; Pred. No. 6.5e+02;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 WNNQOTQFL 15
| | | | | : :
Db 8 WNNQOTQFL 17

RESULT 10
US-10-029-386-31498
Sequence 31498, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
FILE REFERENCE: ABOMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 31498
LENGTH: 20
TYPE: PRT

ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC020728.4
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.99
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.89
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
US-10-029-386-31498

Query Match 31.3% Score 31; DB 14; Length 20;
Best Local Similarity 60.0%; Pred. No. 6.9e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 7 LNNGTQFLF 16
Db 6 NYNQNLIME 15

RESULT 11
US-09-805-290A-15
Sequence 15, Application US/09805290A
Publication No. US20040002583A1
GENERAL INFORMATION:
APPLICANT: BEZEMER, Sandra
APPLICANT: VAN DER BURG, Monique
APPLICANT: DE HAARD, Johannes J
APPLICANT: TAREILUS, Erwin
TITLE OF INVENTION: USE OF ANTIBODIES
FILE REFERENCE: F7526-US[VL,PEK/JVT]seq1a29Aug2001
CURRENT APPLICATION NUMBER: US/09/805,290A
CURRENT FILING DATE: 2001-03-13
PRIOR APPLICATION NUMBER: EP 00200930.6
PRIOR FILING DATE: 2000-03-14
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 15
LENGTH: 12
TYPE: PRT
ORGANISM: lama sp.
US-09-805-290A-15

Query Match 30.3% Score 30; DB 11; Length 12;
Best Local Similarity 62.5%; Pred. No. 5.8e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SSDGLMNN 8
Db 4 SYDGSWMD 11

RESULT 12
US-10-212-499-6
Sequence 6, Application US/10212499
Publication No. US20030135036A1
GENERAL INFORMATION:
APPLICANT: Lanza, Francois
APPLICANT: Phillips, David R.
APPLICANT: Cazenave, Jean-Pierre
TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan Lewis & Bockius LLP
STREET: 1800 M St., NW
CITY: Washington
STATE: DC
COUNTRY: US
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/212,499
FILING DATE: 06-Aug-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/560,814
FILING DATE: 2000-04-28
APPLICATION NUMBER: US 08/089,455
FILING DATE: 1993-07-09
APPLICATION NUMBER: US 08/195,006
FILING DATE: 1994-02-10
APPLICATION NUMBER: US 08/884,571
FILING DATE: 1997-06-27
ATTORNEY/AGENT INFORMATION:
NAME: Reid G. Adler
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 44481-5018-04-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-467-7000
TELEFAX: 202-467-7176
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..13
OTHER INFORMATION: /note="K5/6 peptide residues 1-13."
US-10-212-499-6

Query Match 30.3% Score 30; DB 14; Length 13;
Best Local Similarity 71.4%; Pred. No. 6.3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 12 QLFLEHS 18
Db 7 QLFLEHN 13

RESULT 13
US-09-932-613-254
Sequence 254, Application US/09932613
Publication No. US20030091565A1
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
APPLICANT: Beltzer, James P.
APPLICANT: Potter, M. Daniel
APPLICANT: Fleming, Tony J.
APPLICANT: Rosen, Craig A.
TITLE OF INVENTION: BINDING POLYPEPTIDES AND METHODS BASED THEREON
FILE REFERENCE: DYX-025.1 PCT; DYX-025.1 US
CURRENT APPLICATION NUMBER: US/09/932,613
CURRENT FILING DATE: 2001-08-17
NUMBER OF SEQ ID NOS: 458
SOFTWARE: PatentIn version 3.1
SEQ ID NO 254
LENGTH: 14
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Blys binding polypeptide
US-09-932-613-254

Query Match 30.3% Score 30; DB 10; Length 14;
Best Local Similarity 38.5%; Pred. No. 6.8e+02;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 5 LMNNQTLFLEH 17
Db 2 IWDPLTLMPLH 14

RESULT 14
 US-09-932-322-254
 : Sequence 254, Application US/09932322
 : Publication No. US20030194743A1
 : GENERAL INFORMATION:
 : APPLICANT: Dyax Corp.
 : APPLICANT: Beltzer, James P.
 : APPLICANT: Potter, M. Daniel
 : APPLICANT: Fleming, Tony J.
 : APPLICANT: Ladner, Robert Charles
 : TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (BLYS)
 : FILE REFERENCE: Dvx-018.1 PCT; Dvx-018.1 US
 : CURRENT APPLICATION NUMBER: US/09/932,322
 : CURRENT FILING DATE: 2001-08-17
 : NUMBER OF SEQ ID NOS: 458
 : SOFTWARE: PatentIn version 3.1
 : SEQ ID NO 254
 : LENGTH: 14
 : TYPE: PRT
 : ORGANISM: Artificial Sequence
 : FEATURE:
 : OTHER INFORMATION: Blys binding polypeptide
 : US-09-932-322-254

Query Match	30.3%	Score 30	DB 10	Length 14
Best Local Similarity	38.5%	Pred. No. 6.8e+02		
Matches	5	Conservative	4	Indels 0
				Gaps 0

Qy	5	1	M	N	N	O	T	O	L	E	L	H	17
	:	:	:	:	:	:	:	:	:	:	:	:	
Db	2	I	W	I	D	P	L	T	K	L	M	L	14

```

RESULT 15
US-10-460-512-25
Sequence 25, Application US/10460512
Publication No. US20040038271A1
GENERAL INFORMATION:
APPLICANT: Daniel E.H. Afar
APPLICANT: Rene S. Hubert
APPLICANT: Aya Jakobovits
APPLICANT: Arthur B. Raitano
TITLE OF INVENTION: NOVEL C-TYPE LECTIN TRANSMEMBRANE
FILE REFERENCE: 129.20USU1
CURRENT APPLICATION NUMBER: US/10/460,512
CURRENT FILING DATE: 2003-06-11
PRIOR APPLICATION NUMBER: US/09/638,203
PRIOR FILING DATE: 2000-08-11
PRIOR APPLICATION NUMBER: 60/148,935
PRIOR FILING DATE: 1999-08-12
NUMBER OF SEQ ID NOS: 47
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 25
LENGTH: 14
TYPE: prt
ORGANISM: Homo Sapiens
US-10-460-512-25

```

Query Match	30.3%	Score 30;	DB 15;	Length 14;
Best Local Similarity	55.6%	Pred. No. 6.8e+02;		
Matches	5;	Conservative	4;	Indels 0; Gaps 0;

Oy	4	GLWNNNQ	12
Db	1	GLWRNGDQ	9

RESULT 16
US-10-466-085A-16
; Sequence 16, Application US/10466085A

```

: Publication No. US20040161835A1
: GENERAL INFORMATION:
: APPLICANT: KUBOTA, Michio
: APPLICANT: MARUTA, Kazuhiko
: APPLICANT: YAMAMOTO, Takuo
: APPLICANT: FUKUDA, Shigeharu
: TITLE OF INVENTION: POLYPEPTIDE HAVING -ISOMALTOSYLGLUCOSACCHARIDE-FORMING ENZYME
: TITLE OF INVENTION: ACTIVITY
: FILE REFERENCE: KUBOTA-14
: CURRENT APPLICATION NUMBER: US/10/466,085A
: CURRENT FILING DATE: 2003-07-11
: PRIOR APPLICATION NUMBER: JP 5441/2001
: PRIOR FILING DATE: 2001-01-12
: PRIOR APPLICATION NUMBER: PCT/JP02/00052
: PRIOR FILING DATE: 2001-01-09
: NUMBER OF SEQ ID NOS: 32
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 16
: LENGTH: 17
: TYPE: PRT
: ORGANISM: Bacillus globisporus C11
US-10-466-085A-16

```

```
Query Match      30.3%; Score 30; DB 16; length 17;
Best Local Similarity 57.1%; Pred. No. 8.3e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

Qy	6	WNNNQTO	12
	:	:	
Db	10	WDTNQTE	16

```

RESULT 17
US-09-881-710-15
; Sequence 15, Application US/09881710
; Publication No. US20020086403A1
; GENERAL INFORMATION:
; APPLICANT: DESPRES, Philippe
; APPLICANT: COURAGEOT, Marie-Pierre
; APPLICANT: DEUBEL, Vincent
; APPLICANT: CATTEAU, Adeline
; TITLE OF INVENTION: PRO-APOPTOTIC FRAGMENTS OF THE DENGUE VIRUS ENVELOPE GLYCOPROTEIN
; FILE REFERENCE: 209671050
; CURRENT APPLICATION NUMBER: US/09/881,710
; CURRENT FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/212,129
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Dengue virus
US-09-881-710-15

```

Query Match	30.3%	Score 30	DB 9	Length 20
Best Local	50.0%	Pred. NO	9.9e+02	
Matches	5	Conservative	2	Mismatches 3
				Indels 0
				Gaps 0

Qy	1	SSDGLMNNQ	10
		:	:
		:	:
Db	10	SSGAWKHAQ	19

RESULT 18
US-09-881-710-17

Publication NO. US20020086403A1
GENERAL INFORMATION:
APPLICANT: DESPRES, Philippe

;
; APPLICANT: COURAGEOT, Marie-Pierre
;
; APPLICANT: DEUBEL, Vincent
;
; APPLICANT: CATTEAU, Adeline

TITLE OF INVENTION: PRO-APOPTOTIC FRAGMENTS OF THE DENGUE VIRUS ENVELOPE GLYCOPROTEIN
FILE REFERENCE: 209671US0
CURRENT APPLICATION NUMBER: US/09/881,710
CURRENT FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 60/212,129
PRIOR FILING DATE: 2000-06-16
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 17
LENGTH: 20
TYPE: PRT
ORGANISM: Dengue virus
US-09-881-710-17

Query Match 30.3%; Score 30; DB 9; Length 20;
Best Local Similarity 50.0%; Pred. No. 9.9e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSDGLMNNQ 10
Db 10 SSGAMKHAQ 19

RESULT 19
US-09-973-025-73
Sequence 73, Application US/09973025
Publication No. US20020182706A1
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
BOSMAN, FONS
DE MARTYNOFF, GUY
BUYSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESSES:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/973,025
FILING DATE: 10-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 73:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 73:
US-09-973-025-73

Query Match 30.3%; Score 30; DB 9; Length 20;
Best Local Similarity 38.5%; Pred. No. 9.9e+02;

Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSDGLMNNQTL 13
Db 7 NTNGCMHINSTAL 19

RESULT 20
US-09-973-025-82
Sequence 82, Application US/09973025
Publication No. US20020182706A1
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
BOSMAN, FONS
DE MARTYNOFF, GUY
BUYSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESSES:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/973,025
FILING DATE: 10-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 82:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 82:
US-09-973-025-82

Query Match 30.3%; Score 30; DB 9; Length 20;
Best Local Similarity 40.0%; Pred. No. 9.9e+02;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 WNNQTLFL 15
Db 7 WGENDTLVFV 16

RESULT 21
US-09-899-303-73
Sequence 73, Application US/09899303
Publication No. US20030036110A1
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
BOSMAN, FONS
DE MARTYNOFF, GUY
BUYSE, MARIE-ANGE

TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P. C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,303
FILING DATE: 06-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 73:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 73:
US-09-899-303-73
Query Match 30.3%; Score 30; DB 10; Length 20;
Best Local Similarity 38.5%; Pred. No. 9.9e+02;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
Cy 1 SSDGLMNNQTOFL 13
Db 7 NTNGQWHINSTAL 19
RESULT 22
US-09-899-303-82
Sequence 82, Application US/09899303
Publication No. US20030036110A1
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
BOSMAN, FONS
DE MARTYNOFF, GUY
BUYSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P. C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/899,303
FILING DATE: 06-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 82:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 82:
US-09-899-303-82
Query Match 30.3%; Score 30; DB 10; Length 20;
Best Local Similarity 40.0%; Pred. No. 9.9e+02;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Cy 6 WNNQTOFL 15
Db 7 WGENIDVFLV 16
RESULT 23
US-09-995-808-73
Sequence 73, Application US/09995808
Publication No. US20030095980A1
GENERAL INFORMATION:
APPLICANT: Imogenetics N.V.
TITLE OF INVENTION: Purified hepatitis C virus envelope proteins for diagnostic and
therapeutic use.
FILE REFERENCE: 2551-70
CURRENT APPLICATION NUMBER: US/09/995,808
NUMBER OF SEQ ID NOS: 122
SOFTWARE: Patentin 3.1
SEQ ID NO 73
LENGTH: 20
TYPE: PRT
ORGANISM: Hepatitis C virus
US-09-995-808-73
Query Match 30.3%; Score 30; DB 10; Length 20;
Best Local Similarity 38.5%; Pred. No. 9.9e+02;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
Cy 1 SSDGLMNNQTOFL 13
Db 7 NTNGQWHINSTAL 19
RESULT 24
US-09-995-808-82
Sequence 82, Application US/09995808
Publication No. US20030095980A1
GENERAL INFORMATION:
APPLICANT: Imogenetics N.V.
TITLE OF INVENTION: Purified hepatitis C virus envelope proteins for diagnostic and
therapeutic use.
FILE REFERENCE: 2551-70
CURRENT APPLICATION NUMBER: US/09/995,808
NUMBER OF SEQ ID NOS: 122
SOFTWARE: Patentin 3.1

SEQ ID NO 82
LENGTH: 20
TYPE: PRT
ORGANISM: Hepatitis C virus
US-09-995-808-82

Query Match 30.3%; Score 30; DB 10; Length 20;
Best Local Similarity 40.0%; Pred. No. 9.9e+02;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 WNNQTOLEFL 15
DB 7 WGENDTDFV 16

RESULT 25
US-09-995-860-73
Sequence 73, Application US/09995860
Publication No. US20030118603A1
GENERAL INFORMATION:

APPLICANT: Innogenetics N.V.
TITLE OF INVENTION: Purified hepatitis C virus envelope proteins for diagnostic and
FILE REFERENCE: 2551-69
CURRENT APPLICATION NUMBER: US/09/995,860
CURRENT FILING DATE: 2001-11-29
NUMBER OF SEQ ID NOS: 122
SOFTWARE: PatentIn 3.1
SEQ ID NO 73
LENGTH: 20
TYPE: PRT
ORGANISM: Hepatitis C virus
US-09-995-860-73

Query Match 30.3%; Score 30; DB 10; Length 20;
Best Local Similarity 38.5%; Pred. No. 9.9e+02;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSDGLMNNQTOQL 13
DB 7 NTNGQWHINSTAL 19

RESULT 26
US-09-995-860-82

Sequence 82, Application US/09995860
Publication No. US20030118603A1
GENERAL INFORMATION:
APPLICANT: Innogenetics N.V.
TITLE OF INVENTION: Purified hepatitis C virus envelope proteins for diagnostic and
FILE REFERENCE: 2551-69
CURRENT APPLICATION NUMBER: US/09/995,860
CURRENT FILING DATE: 2001-11-29
NUMBER OF SEQ ID NOS: 122
SOFTWARE: PatentIn 3.1
SEQ ID NO 82
LENGTH: 20
TYPE: PRT
ORGANISM: Hepatitis C virus
US-09-995-860-82

Query Match 30.3%; Score 30; DB 10; Length 20;
Best Local Similarity 40.0%; Pred. No. 9.9e+02;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 WNNQTOLEFL 15
DB 7 WGENDTDFV 16

RESULT 27
US-09-995-791-73

Sequence 73, Application US/09995791
Publication No. US20030147918A1
GENERAL INFORMATION:

APPLICANT: Innogenetics N.V.
TITLE OF INVENTION: Purified hepatitis C virus envelope proteins for diagnostic and
FILE REFERENCE: 2551-68
CURRENT APPLICATION NUMBER: US/09/995,791
CURRENT FILING DATE: 2001-11-29
NUMBER OF SEQ ID NOS: 122
SOFTWARE: PatentIn 3.1
SEQ ID NO 73
LENGTH: 20
TYPE: PRT
ORGANISM: Hepatitis C virus
US-09-995-791-73

Query Match 30.3%; Score 30; DB 10; Length 20;
Best Local Similarity 38.5%; Pred. No. 9.9e+02;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSDGLMNNQTOQL 13
DB 7 NTNGQWHINSTAL 19

RESULT 28
US-09-995-791-82
Sequence 82, Application US/09995791
Publication No. US20030147918A1
GENERAL INFORMATION:

APPLICANT: Innogenetics N.V.
TITLE OF INVENTION: Purified hepatitis C virus envelope proteins for diagnostic and
FILE REFERENCE: 2551-68
CURRENT APPLICATION NUMBER: US/09/995,791
CURRENT FILING DATE: 2001-11-29
NUMBER OF SEQ ID NOS: 122
SOFTWARE: PatentIn 3.1
SEQ ID NO 82
LENGTH: 20
TYPE: PRT
ORGANISM: Hepatitis C virus
US-09-995-791-82

Query Match 30.3%; Score 30; DB 10; Length 20;
Best Local Similarity 40.0%; Pred. No. 9.9e+02;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 WNNQTOLEFL 15
DB 7 WGENDTDFV 16

RESULT 29
US-10-651-165-38

Sequence 38, Application US/10651165
Publication No. US2004004877A1
GENERAL INFORMATION:
APPLICANT: LEROUX-ROELS, Geert
APPLICANT: DELEYS, Robert
APPLICANT: MAERTENS, Geert
TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
FILE REFERENCE: 2551-94
CURRENT APPLICATION NUMBER: US/10/651,165
CURRENT FILING DATE: 2003-09-02
PRIOR APPLICATION NUMBER: US/08/974,690C
PRIOR FILING DATE: 1997-11-19
PRIOR APPLICATION NUMBER: PCT/EP94/03555
PRIOR FILING DATE: 1994-10-28
PRIOR APPLICATION NUMBER: EP 93402718.6
PRIOR FILING DATE: 1993-11-04

```
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38
; LENGTH: 20
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-10-651-165-38
```

```
Query Match          30.3%; Score 30; DB 15; Length 20;
Best Local Similarity 40.0%; Pred. No. 9.9e+02;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
QY      6 WNNNQTLFL 15
      |||||
Db      7 WGENDTDFEV 16
```

```
RESULT 30
US-10-651-165-47
; Sequence 47, Application US/10651165
; Publication No. US20040047877A1
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELAIS, Robert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/10/651,165
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US/08/974,690C
; PRIOR FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47
; LENGTH: 20
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-10-651-165-47
```

```
Query Match          30.3%; Score 30; DB 15; Length 20;
Best Local Similarity 38.5%; Pred. No. 9.9e+02;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
```

```
QY      1 SSDGLMNNQTL 13
      ::|||
Db      7 NTNGSMHNSYAL 19
```

```
RESULT 31
US-10-651-165-56
; Sequence 56, Application US/10651165
; Publication No. US20040047877A1
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELAIS, Robert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/10/651,165
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US/08/974,690C
; PRIOR FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
```

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56
; LENGTH: 20
; TYPE: PRT
; ORGANISM: hepatitis C virus
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (2)..(2)
; OTHER INFORMATION: Xaa is Asn, Lys, Asp or Arg
; NAME/KEY: MISC_FEATURE
; LOCATION: (3)..(3)
; OTHER INFORMATION: Xaa is Val, Ile or Leu
; NAME/KEY: MISC_FEATURE
; LOCATION: (4)..(4)
; OTHER INFORMATION: Xaa is Gln, Ser or Tyr
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (6)..(6)
; OTHER INFORMATION: Xaa is Ile or Val
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (8)..(8)
; OTHER INFORMATION: Xaa is Thr or Ser
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (14)..(14)
; OTHER INFORMATION: Xaa is Leu or Ile
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (16)..(16)
; OTHER INFORMATION: Xaa is Ser or Arg
US-10-651-165-56
```

```
Query Match          30.3%; Score 30; DB 15; Length 20;
Best Local Similarity 38.5%; Pred. No. 9.9e+02;
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 SSDGLMNNQTL 13
      ::|||
Db      7 NXNGSMHNSYAL 19
```

```
RESULT 32
US-10-634-895-15
; Sequence 15, Application US/10634895
; Publication No. US20040049016A1
; GENERAL INFORMATION:
; APPLICANT: DESPRES, Philippe
; APPLICANT: COURAGEOT, Marie-Pierre
; APPLICANT: DEUBEL, Vincent
; APPLICANT: CATTEAU, Adeline
; TITLE OF INVENTION: PRO-APOTOTIC FRAGMENTS OF THE DENGUE VIRUS ENVELOPE GLYCOPROTEIN
; FILE REFERENCE: 209671U50
; CURRENT APPLICATION NUMBER: US/10/634,895
; CURRENT FILING DATE: 2003-08-06
; PRIOR APPLICATION NUMBER: US/09/881,710
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/212,129
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Dengue virus
US-10-634-895-15
```

```
Query Match          30.3%; Score 30; DB 15; Length 20;
Best Local Similarity 50.0%; Pred. No. 9.9e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

OY 1 SSDGLMNNQ 10
||:|:|:|
Db 10 SSEGAMKXIAQ 19

RESULT 33
US-10-634-895-17
; Sequence 17, Application US/10634895
; Publication No. US20040049016A1
; GENERAL INFORMATION:
; APPLICANT: DESPRES, Philippe
; APPLICANT: COURAGEOT, Marie-Pierre
; APPLICANT: DEUBEL, Vincent
; APPLICANT: CARTEAU, Adeline
; TITLE OF INVENTION: PRO-APOPTOTIC FRAGMENTS OF THE DENGUE VIRUS ENVELOPE GLYCOPROTEIN
; FILE REFERENCE: 209671USO
; CURRENT APPLICATION NUMBER: US/10/634,895
; CURRENT FILING DATE: 2003-08-06
; PRIOR APPLICATION NUMBER: US/09/881,710
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/212,129
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Dengue virus
US-10-634-895-17

Query Match 30.3%; Score 30; DB 15; Length 20;
Best Local Similarity 50.0%; Pred. No. 9.9e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 SSDGLMNNQ 10
||:|:|:|
Db 10 SSEGAMKXIAQ 19

RESULT 34
US-10-311-213-22
; Sequence 22, Application US/10311213
; Publication No. US20040101862A1
; GENERAL INFORMATION:
; APPLICANT: DESPRES, Philippe
; APPLICANT: COURAGEOT, Marie-Pierre
; APPLICANT: DEUBEL, Vincent
; APPLICANT: CARTEAU, Adeline
; TITLE OF INVENTION: PRO-APOPTOTIC FRAGMENTS OF THE DENGUE VIRUS ENVELOPE GLYCOPROTEIN
; FILE REFERENCE: 231164USOPT
; CURRENT APPLICATION NUMBER: US/10/311,213
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/IB01/01570
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/212,129
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Dengue virus type 4
US-10-311-213-22

Query Match 30.3%; Score 30; DB 16; Length 20;
Best Local Similarity 50.0%; Pred. No. 9.9e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 SSDGLMNNQ 10
||:|:|:|
Db 10 SSEGAMKXIAQ 19

RESULT 35
US-10-321-798-73
; Sequence 73, Application US/10321798
; Publication No. US20040126395A1
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: Purified hepatitis C virus envelope proteins for diagnostic and
; FILE REFERENCE: 2551-93
; CURRENT APPLICATION NUMBER: US/10/321,798
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: 60/418,358
; PRIOR FILING DATE: 2002-10-16
; PRIOR APPLICATION NUMBER: 10/020,510
; PRIOR FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn 3.1
; SEQ ID NO 73
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-321-798-73

Query Match 30.3%; Score 30; DB 16; Length 20;
Best Local Similarity 38.5%; Pred. No. 9.9e+02;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 1 SSDGLMNNQ 13
||:|:|:|
Db 7 NTNGQWHINSTAL 19

RESULT 36
US-10-321-798-82
; Sequence 82, Application US/10321798
; Publication No. US20040126395A1
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: Purified hepatitis C virus envelope proteins for diagnostic and
; FILE REFERENCE: 2551-93
; CURRENT APPLICATION NUMBER: US/10/321,798
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: 60/418,358
; PRIOR FILING DATE: 2002-10-16
; PRIOR APPLICATION NUMBER: 10/020,510
; PRIOR FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn 3.1
; SEQ ID NO 82
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-321-798-82

Query Match 30.3%; Score 30; DB 16; Length 20;
Best Local Similarity 40.0%; Pred. No. 9.9e+02;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 6 WNNNQTOFL 15
||:|:|:|
Db 7 WGENDTDFV 16

RESULT 37
US-09-563-222-53
; Sequence 53, Application US/09563222
; Publication No. US20030079253A1
; GENERAL INFORMATION:
; APPLICANT: Hiact, Andrew
; APPLICANT: Hehn, Mich B.
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN ARRAYS IN
; EUKARYOTIC CELLS

FILE REFERENCE: 310098.406
CURRENT APPLICATION NUMBER: US/09/563.222
CURRENT FILING DATE: 2000-05-02
NUMBER OF SEQ ID NOS: 197
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 53
LENGTH: 7
TYPE: PRT
ORGANISM: Mus musculus
US-09-563-222-53

Query Match 29.3% Score 29; DB 10; Length 7;
Best Local Similarity 66.7% Pred. No. 1.5e+06;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 GLMNNN 9
| | | | |
Db 2 GYMNS 7

RESULT 38
US-10-783-950-53
Sequence 53, Application US/10783950
Publication No. US20040199945A1
GENERAL INFORMATION:
APPLICANT: EPICYTE PHARMACEUTICALS, INC.
APPLICANT: HEIN, MICHAEL B.
TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN ARRAYS IN PLANT CELLS
FILE REFERENCE: 068904-0501
CURRENT APPLICATION NUMBER: US/10/783.950
CURRENT FILING DATE: 2004-02-19
PRIOR APPLICATION NUMBER: US/09/563.222
PRIOR FILING DATE: 2000-05-02
PRIOR APPLICATION NUMBER: PCT/US01/14349
PRIOR FILING DATE: 2001-05-02
PRIOR APPLICATION NUMBER: 09/563,222
PRIOR FILING DATE: 2000-05-02
NUMBER OF SEQ ID NOS: 182
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 53
LENGTH: 7
TYPE: PRT
ORGANISM: Mus musculus
US-10-783-950-53

Query Match 29.3% Score 29; DB 17; Length 7;
Best Local Similarity 66.7% Pred. No. 1.5e+06;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 GLMNNN 9
| | | | |
Db 2 GYMNS 7

RESULT 39
US-10-416-249-557
Sequence 557, Application US/10416249
Publication No. US20040132121A1
GENERAL INFORMATION:
APPLICANT: Dairymple, Brian P.
APPLICANT: Kongsuwan, Kitaya
APPLICANT: Wiffels, Gene L.
APPLICANT: Jennings, Philip A.
TITLE OF INVENTION: METHOD OF IDENTIFYING ANTIBACTERIAL
FILE REFERENCE: CULN42.001APC
CURRENT APPLICATION NUMBER: US/10/416.249
CURRENT FILING DATE: 2003-11-10
PRIOR APPLICATION NUMBER: PCT/AU01/01436
PRIOR FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: AU PR 1320

PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: AU PR 2919
PRIOR FILING DATE: 2001-02-06
NUMBER OF SEQ ID NOS: 678
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 557
LENGTH: 15
TYPE: PRT
ORGANISM: Bacillus halodurans
US-10-416-249-557

Query Match 29.3% Score 29; DB 16; Length 15;
Best Local Similarity 50.0% Pred. No. 1.1e+03;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 DGLMNNNOTQLF 14
: | | : | | | |
Db 4 EGLPDSNQSLF 15

RESULT 40
US-10-416-249-563
Sequence 563, Application US/10416249
Publication No. US20040132121A1
GENERAL INFORMATION:
APPLICANT: Dairymple, Brian P.
APPLICANT: Kongsuwan, Kitaya
APPLICANT: Wiffels, Gene L.
APPLICANT: Jennings, Philip A.
TITLE OF INVENTION: METHOD OF IDENTIFYING ANTIBACTERIAL
FILE REFERENCE: CULN42.001APC
CURRENT APPLICATION NUMBER: US/10/416.249
CURRENT FILING DATE: 2003-11-10
PRIOR APPLICATION NUMBER: PCT/AU01/01436
PRIOR FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: AU PR 1320
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: AU PR 2919
PRIOR FILING DATE: 2001-02-06
NUMBER OF SEQ ID NOS: 678
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 563
LENGTH: 15
TYPE: PRT
ORGANISM: Listeria innocua
US-10-416-249-563

Query Match 29.3% Score 29; DB 16; Length 15;
Best Local Similarity 50.0% Pred. No. 1.1e+03;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 DGLMNNNOTQLF 14
: | | : | | | |
Db 4 EGLPDSNQSLF 15

Search completed: December 30, 2004, 16:43:26
Job time: 144 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 16:22:44 ; Search time 38 Seconds
(without alignments)
45.576 Million cell updates/sec

Title: US-10-718-321-1

Perfect score: 99

Sequence: 1 SSDLMNNNGTQLFLHS 18

Scoring table: BLOSUM62

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 3886

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%

Listing first 150 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31.3	16	2	PA9039	T-cell receptor be
2	30.3	19	2	I49422	L-lactate dehydrog
3	27.3	20	2	PQ0300	large protein - bo
4	26.3	16	2	PH0749	T-cell receptor be
5	26.3	19	2	I46654	T-cell receptor de
6	25.3	13	2	PH0138	T-cell receptor be
7	24.2	6	2	PT0589	T-cell receptor be
8	24.2	10	2	S23370	T-cell receptor al
9	24.2	13	2	PH0756	T-cell receptor be
10	24.2	16	2	PC1299	subtilisin (EC 3.4
11	24.2	20	2	S27351	lysophospholipase
12	23.7	20	2	S38763	S-adenosyl-L-methi
13	23.2	5	2	UH0253	gut pentapeptide -
14	23.2	8	2	CE1512	variant surface gl
15	23.2	8	2	A41117	acetylcholinestera
16	23.2	10	2	PH1344	Ig heavy chain DJ
17	23.2	10	2	A44874	proboscipedia - fr
18	23.2	15	2	S51735	T-cell receptor be
19	23.2	17	2	AE1211	ananthin - Streptom
20	23.2	18	2	CS6211	progesterone recep
21	23.2	19	2	A44854	L-2,4-diaminobuty
22	23.2	20	2	S16202	pyrroline-5-carbox
23	23.2	20	2	S56756	link protein - rat
24	23.2	10	2	PH0923	T-cell receptor be
25	22.2	14	2	PH1471	T-cell receptor be
26	22.2	15	2	G49655	T-cell receptor be
27	22.2	15	2	B45474	chromospondin 2 -
28	22.2	17	2	S57519	T cell receptor be
29	22.2	17	2	S57556	T cell receptor be

30	22.2	18	2	I40062	shikimate 5-dehydr
31	22.2	19	2	PC1251	teatin II - rat (f
32	22.2	19	2	I40063	shikimate 5-dehydr
33	22.2	20	2	H22565	R-phycocerythrin ga
34	21.2	8	2	A39308	glycine reductase
35	21.2	9	2	PT0562	T-cell receptor be
36	21.2	12	2	PH1308	Ig heavy chain DJ
37	21.2	14	2	S57572	T cell receptor V-
38	21.2	14	2	PH1594	Ig H chain V-D-J r
39	21.2	16	2	A59046	alpha-conotoxin MI
40	21.2	16	2	S28213	glutathione transf
41	21.2	16	2	PH0763	T-cell receptor be
42	21.2	16	2	PH1474	T-cell receptor be
43	21.2	18	2	S43834	DNA topoisomerase
44	21.2	18	2	I46653	T-cell receptor de
45	21.2	19	2	A28814	Ig kappa chain V r
46	21.2	19	2	I52721	gene hMLH1 protein
47	21.2	20	2	B60505	hemoglobin A1-2 be
48	20.2	5	2	PT0659	T-cell receptor be
49	20.2	6	2	PT0657	T-cell receptor be
50	20.2	6	2	PT0720	T-cell receptor be
51	20.2	7	2	PT0611	T-cell receptor be
52	20.2	7	2	PT0523	T-cell receptor be
53	20.2	7	2	PT0529	T-cell receptor be
54	20.2	8	2	PT0613	T-cell receptor be
55	20.2	8	2	PT0725	T-cell receptor be
56	20.2	10	2	PH0925	T-cell receptor be
57	20.2	15	2	PH0008	lectin B2 - Psopho
58	20.2	16	2	PH1472	T-cell receptor be
59	20.2	16	2	A49226	major outer membra
60	20.2	16	2	C37290	homeotic protein G
61	20.2	17	2	C43599	hypothetical prote
62	20.2	17	2	S57555	T cell receptor V-
63	20.2	17	2	B31769	T-cell receptor de
64	20.2	17	2	PH0757	T-cell receptor be
65	20.2	17	2	A27636	cytochrome B - C10s
66	20.2	19	2	A28702	T-cell receptor de
67	20.2	20	2	I46652	T-cell receptor al
68	20.2	20	2	S03505	T-cell receptor al
69	19.2	6	2	PT0532	T-cell receptor be
70	19.2	7	2	S08606	hypothetical prote
71	19.2	9	2	I52974	seminal vesicle pr
72	19.2	10	2	S59625	beta-galactosidase
73	19.2	10	2	I40032	type protein - Bac
74	19.2	12	2	S25039	Ig heavy chain V r
75	19.2	12	2	PH1470	T-cell receptor be
76	19.2	13	2	PT0293	Ig heavy chain CRD
77	19.2	13	2	S57567	T cell receptor V-
78	19.2	14	2	PH1757	T cell receptor al
79	19.2	14	2	B44854	L-2,4-diaminobuty
80	19.2	15	2	PH1788	T cell receptor al
81	19.2	15	2	D46743	corneal keratan su
82	19.2	15	2	PQ0073	T-cell receptor be
83	19.2	16	2	F53284	T-cell receptor be
84	19.2	16	2	B44896	heat shock protein
85	19.2	16	2	E37290	homeotic protein G
86	19.2	17	2	S51736	T-cell receptor be
87	19.2	17	2	S77834	DNA-directed RNA p
88	19.2	19	2	EX0062	beta-galactoside-b
89	19.2	20	2	PC2347	base nonpecific a
90	19.2	20	2	PM0133	pepsin (EC 3.4.23.
91	19.2	20	2	B32384	G-alpha-2 protein
92	19.2	20	2	A44773	pollen allergen I
93	19.2	20	2	A34817	collagenolytic pro
94	19.2	20	2	A34859	heletherrhine - Mex
95	19.2	20	2	A41439	acid ribonuclease
96	18.5	11	2	PH0904	T-cell receptor be
97	18.7	20	2	PC1240	calcium-binding pr
98	18.2	7	1	NYR37	hypotatamic hepta
99	18.2	7	2	PT0521	T-cell receptor be
100	18.2	8	2	PT0716	T-cell receptor be
101	18.2	9	2	E28854	fibrinopeptide B -
102	18.2	9	2	F28854	fibrinopeptide B -

103	18	18.2	9	2	D28854	fibrinopeptide B -
104	18	18.2	9	2	PT0238	Ig heavy chain CRD
105	18	18.2	9	2	PT0272	Ig heavy chain CRD
106	18	18.2	10	2	P00753	beta-fructofuranos
107	18	18.2	11	2	PH0914	T-cell receptor be
108	18	18.2	11	2	S04875	nifs protein - Bra
109	18	18.2	12	2	A29169	phospholipase A2 (
110	18	18.2	12	2	S68271	major urinary prot
111	18	18.2	12	2	PT0216	T-cell receptor be
112	18	18.2	12	2	S28215	glucan endo-1,3-be
113	18	18.2	12	2	A49637	MHC class II histo
114	18	18.2	13	2	S47358	T-cell antigen rec
115	18	18.2	14	2	PT0077	proteochondotin c
116	18	18.2	14	2	S17766	beta-glucosidase (
117	18	18.2	14	2	PA0007	lectin B1 - Psopho
118	18	18.2	15	1	SPRT	scotophobin - rat
119	18	18.2	15	2	I46512	tropomn - rabbit
120	18	18.2	15	2	PH1366	Ig heavy chain DJ
121	18	18.2	15	2	PH0782	T-cell receptor al
122	18	18.2	16	2	A46236	transforming prote
123	18	18.2	17	2	I46511	tyroponin - rabbit
124	18	18.2	17	2	E53113	gingipain, 44K - P
125	18	18.2	17	2	B61491	seed protein ws-2
126	18	18.2	17	2	B14335	adherence lectin 1
127	18	18.2	17	2	C94063	hypothetical prote
128	18	18.2	17	4	I76673	hypothetical COI1/
129	18	18.2	18	2	B48839	T-cell receptor be
130	18	18.2	18	2	H64711	hypothetical prote
131	18	18.2	18	2	I50389	mosin heavy chain
132	18	18.2	18	4	I54078	hypothetical PMU/R
133	18	18.2	19	2	P00548	capsid protein VP2
134	18	18.2	19	2	S74087	antibacterial prot
135	18	18.2	19	2	I39327	pre-T/NK cell-grow
136	18	18.2	20	2	PM0115	insulin-like growt
137	18	18.2	20	2	S32502	calpain (BC 3.4.22
138	18	18.2	20	2	PH1326	Ig heavy chain DJ
139	18	18.2	20	2	S66222	defensin AMP2 - Da
140	18	18.2	20	2	B44913	periplasmic flagel
141	18	18.2	20	2	S28435	major outer membra
142	18	18.2	20	2	S03987	agglutinin beta-20
143	18	18.2	20	2	B61497	seed protein ws-20
144	17	17.2	6	2	PT0519	T-cell receptor be
145	17	17.2	7	2	PT0688	T-cell receptor be
146	17	17.2	8	2	PT0724	T-cell receptor be
147	17	17.2	8	2	XGHUEU	urine glycopeptide
148	17	17.2	9	2	C24180	fibrinogen beta ch
149	17	17.2	9	2	A45199	L-hyosoplorin - Ja
150	17	17.2	10	2	S71868	glutathione transf

ALIGNMENTS

RESULT 1
 T-cell receptor beta chain V-D-J-C region (V beta 2, J beta 1.5) - human (fragment)
 C/Species: Homo sapiens (man)
 C/Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 30-May-1997
 C/Accession: F49039
 R/Rosenberg, W.M.; Moss, P.A.; Bell, J.I.
 Eur. J. Immunol. 22, 541-549, 1992
 A/Title: Variation in human T cell receptor V beta and J beta repertoire: analysis using
 A/Reference number: A49039; MUID:92164737; PMID:1311263
 C/Accession: F49039
 A/Status: preliminary; not compared with conceptual translation
 A/Molecule type: nucleic acid
 A/Residues: 1-16 <ROS>
 A/Note: sequence extracted from NCBI backbone (NCBI:90718)
 C/Keywords: T-cell receptor

Query Match 31.3%; Score 31; DB 2; Length 16;
 Best Local Similarity 50.0%; Pred. No. 86;
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SSDGLMNNQTOLEF 14
 Db 2 SASGTDNSNQPHF 15

RESULT 2
 I49422
 L-lactate dehydrogenase (EC 1.1.1.27) chain X - western wild mouse (fragment)
 M/Alternate names: lactate dehydrogenase C
 C/Species: Mus spretus (western wild mouse)
 C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
 C/Accession: I49422
 R/Ko, M.S.; Wang, X.; Horton, J.H.; Hagen, M.D.; Takahashi, N.; Maezaki, Y.; Nadeau, J.H
 Mamm. Genome 5, 349-355, 1994
 A/Title: Genetic mapping of 40 cDNA clones on the mouse genome by PCR.
 A/Reference number: I48934; MUID:94319082; PMID:8043949
 C/Accession: I49422
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-19 <RES>
 A/Cross-references: UNIPROT:Q62545; EMBL:U05739; NID:g497076; PIDN:AAB60478.1; PID:g49707
 C/Superfamily: L-lactate dehydrogenase
 C/Keywords: oxidoreductase

Query Match 30.3%; Score 30; DB 2; Length 19;
 Best Local Similarity 71.4%; Pred. No. 1.5e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SSDGLMNN 7
 Db 6 SADTLTMN 12

RESULT 3
 PQ0300
 large protein - bovine respiratory syncytial virus (strain A51908) (fragment)
 C/Species: bovine respiratory syncytial virus
 C/Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004
 C/Accession: PQ0300
 R/Zamorin, M.; Samal, S.K.
 J. Gen. Virol. 73, 737-741, 1992
 A/Title: Sequence analysis of M2 mRNA of bovine respiratory syncytial virus obtained from
 A/Reference number: JQ1481; MUID:92185490; PMID:1312110
 C/Accession: PQ0300
 A/Molecule type: mRNA
 A/Residues: 1-20 <ZAM>
 A/Cross-references: UNIPROT:Q65705; UNIPROT:Q8V686; GB:M82816; NID:9210823; PIDN:AAA4280
 C/Genetics:
 A/Gene: L

Query Match 27.3%; Score 27; DB 2; Length 20;
 Best Local Similarity 37.5%; Pred. No. 4.8e+02;
 Matches 6; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 3 DGLMNNQTOLEFHS 18
 Db 2 DTLIHSTWVYLTDS 17

RESULT 4
 PH0749
 T-cell receptor beta chain (B83) - mouse (fragment)
 C/Species: Mus musculus (house mouse)
 C/Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 05-Nov-1999
 C/Accession: PH0749
 R/Casanova, J.U.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
 J. Exp. Med. 174, 1371-1383, 1991
 A/Title: T cell receptor genes in a series of class I major histocompatibility complex-r
 allelic exclusion and antigen-specific repertoire.
 A/Reference number: PH0746; MUID:92078846; PMID:1836010
 C/Accession: PH0749
 A/Molecule type: mRNA

A:Residues: 1-16 <CAS>
 A:Cross-references: EMBL:X60840; NID:950116; PIDN:CAA4323.1; PID:950117
 A:Experimental source: T lymphocyte
 C:Keywords: T-cell receptor

Query Match 26.3%; Score 26; DB 2; Length 16;
 Best Local Similarity 71.4%; Pred. No. 5.4e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 NNQOLF 14
 |||||
 Db 9 NNQAPLF 15

RESULT 5

146654 T-cell receptor delta-chain J-delta-3 - pig (fragment)

C:Species: Sus scrofa domestica (domestic pig)
 C>Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 05-Nov-1999

C:Accession: 146654
 R:Yang, Y.G.; Ohta, S.; Yamada, S.; Shimizu, M.; Takagaki, Y.

J. Immunol. 153, 1981-1993, 1995
 A:Title: Diversity of T cell receptor delta-chain cDNA in the thymus of a one-month-old

A:Reference number: 146623; MUID:95363165; PMID:7636249

A:Accession: 146654

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-19 <YAN>

A:Cross-references: GB:D49562; NID:G1041176; PIDN:BAA08506.1; PID:G1041177

C:Keywords: T-cell receptor

Query Match 26.3%; Score 26; DB 2; Length 19;
 Best Local Similarity 35.3%; Pred. No. 6.5e+02;
 Matches 6; Conservative 3; Mismatches 2; Indels 6; Gaps 1;

QY 6 WNNNN-----TOLFLE 16
 |||||
 Db 2 WDTROMYFGACTKLFVE 18

RESULT 6
 PH0138 T-cell receptor beta chain V-D-J region C8 - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 30-May-1997

C:Accession: PH0138
 R:Martin, R.; Howell, M.D.; Jazarequema, D.; Flerlage, M.; Brostoff, S.; Le

J. Exp. Med. 173, 19-24, 1991
 A:Title: A myelin basic protein peptide is recognized by cytotoxic T cells in the context

A:Reference number: PH0135; MUID:91086643; PMID:1702137

A:Accession: PH0138

A:Molecule type: mRNA

A:Residues: 1-13 <MAR>

A:Cross-references: T-cell receptor

C:Keywords: T-cell receptor

Query Match 25.3%; Score 25; DB 2; Length 13;
 Best Local Similarity 60.0%; Pred. No. 6.2e+02;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 WNNNQ 10
 |||||
 Db 7 WTNNE 11

RESULT 7
 PT0589 T-cell receptor beta chain V-D-J region (141-1A) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

A:Reference number: PT0509; MUID:91277601; PMID:1711558
 A:Accession: PT0589
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-6 <PEE>
 A:Experimental source: day 19 fetal thymus, strain BALB/c
 C:Keywords: T-cell receptor

Query Match 24.2%; Score 24; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.6e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSDGL 5
 |||||
 Db 2 SSDGL 6

RESULT 8

S23370 T-cell receptor alpha chain J region - human (fragment)

C:Species: Homo sapiens (man)
 C>Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 17-Mar-1999

C:Accession: S23370

R:Pluschke, G.; Ricken, G.; Taube, H.; Kroninger, S.; Welchers, I.; Peter, H.H.; Eichman

Bur. J. Immunol. 21, 2749-2754, 1991

A:Title: Biased T cell receptor (Valpha) region repertoire in the synovial fluid of rheu

A:Reference number: S23364; MUID:92037820; PMID:1657615

A:Accession: S23370

A:Status: preliminary; translation not shown

A:Molecule type: mRNA

A:Residues: 1-10 <PLU>

A:Cross-references: EMBL:X58165

C:Keywords: T-cell receptor

Query Match 24.2%; Score 24; DB 2; Length 10;
 Best Local Similarity 75.0%; Pred. No. 6.7e+02;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 WNNN 9
 |||||
 Db 3 WDNN 6

RESULT 9
 PH0756 T-cell receptor beta chain (I7) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 05-Nov-1999

C:Accession: PH0756

R:Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.

J. Exp. Med. 174, 1371-1383, 1991

A:Title: T cell receptor genes in a series of class I major histocompatibility complex-r

allelic exclusion and antigen-specific repertoire.

A:Reference number: PH0746; MUID:92078846; PMID:1836010

A:Accession: PH0756

A:Molecule type: mRNA

A:Residues: 1-13 <CAS>

A:Cross-references: EMBL:X60850; NID:951482; PIDN:CAA43241.1; PID:951483

A:Experimental source: T lymphocyte

C:Keywords: T-cell receptor

Query Match 24.2%; Score 24; DB 2; Length 13;
 Best Local Similarity 50.0%; Pred. No. 8.9e+02;
 Matches 7; Conservative 0; Mismatches 3; Indels 4; Gaps 1;

QY 1 SSDGLNNNQOLF 14
 |||||
 Db 3 SGDGL----NQAPLF 12

RESULT 10
 PC1299 subtilisin (BC 3.4.21.62) GX - Bacillus sp. (strain 6644) (fragment)

C/Species: *Bacillus* sp.
C/Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
C/Accession: PC1299
R/Durham, D.R.
Biochem. Biophys. Res. Commun. 194, 1365-1370, 1993
A/Title: The elastolytic properties of subtilisin GX from alkalophilic *Bacillus* sp. str.
A/Reference number: PC1299; PMID:93356814; PMID:8352796
A/Accession: PC1299
A/Molecule type: protein
A/Residues: 1-16 <DUP>
A/Cross-references: UNIPROT:Q9R557
C/Keywords: hydrolase; serine proteinase

Query Match 24.2%; Score 24; DB 2; Length 16;
Best Local Similarity 42.9%; Pred. No. 1.1e+03;
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 6 WNNNQTLQ 12
Db 5 WQDNRVQ 11

RESULT 11
S27351
lysophospholipase - human
C/Species: *Homo sapiens* (man)
C/Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C/Accession: S27351
R/Garsetti, D.; Holtsberg, F.; Steiner, M.R.; Egan, R.W.; Clark, M.A.
Biochem. J. 288, 831-837, 1992
A/Title: Butyric acid-induced differentiation of HU-60 cells increases the expression of
A/Reference number: S27350; PMID:93111958; PMID:1471998
A/Accession: S27351
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-20 <GAK>
A/Cross-references: UNIPROT:P56643

Query Match 24.2%; Score 24; DB 2; Length 20;
Best Local Similarity 35.7%; Pred. No. 1.4e+03;
Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 SSDGLMNNQTLQ 14
Db 2 ASISLYNSNTLSYF 15

RESULT 12
S38763
S-adenosyl-L-methionine decarboxylase alpha chain - *Acanthamoeba castellanii*
C/Species: *Acanthamoeba castellanii*
C/Date: 08-Jun-1994 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
C/Accession: S38763
R/Hugo, E.R.; Byers, T.J.
Biochem. J. 293, 203-209, 1993
A/Title: S-Adenosyl-L-methionine decarboxylase of *Acanthamoeba castellanii* (Neff): purifi
A/Reference number: S38763; PMID:94029912; PMID:8216217
A/Accession: S38763
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-20 <HUG>
A/Cross-references: UNIPROT:P34039

Query Match 23.7%; Score 23.5; DB 2; Length 20;
Best Local Similarity 43.8%; Pred. No. 1.7e+03;
Matches 7; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

Qy 1 SSDGLMNNQTLQ 16
Db 2 SSMFVNM---TKLILX 14

RESULT 13

JH0253
gut pentapeptide - Japanese eel
C/Species: *Anguilla japonica* (Japanese eel)
C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 11-Apr-1995
C/Accession: JH0253
R/Uesaka, T.; Ikeda, T.; Kubota, I.; Munesaka, Y.; Ando, M.
Biochem. Biophys. Res. Commun. 180, 828-832, 1991
A/Title: Structure and function of a pentapeptide isolated from the gut of the eel.
A/Reference number: JH0253; PMID:92062113; PMID:1953755
A/Accession: JH0253
A/Molecule type: protein
A/Residues: 1-5 <UES>
A/Experimental source: gut
C/Comment: This peptide increased basal tone of the circular muscle of the esophagogastric
, and of the circular muscle of the gastro-intestinal junction.

Query Match 23.2%; Score 23; DB 2; Length 5;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GLMN 7
Db 1 GFNM 4

RESULT 14
C61512
variant surface glycoprotein MITac 1.4 - *Trypanosoma brucei* (fragment)
C/Species: *Trypanosoma brucei*
C/Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jul-2004
C/Accession: C61512
R/Holzer, A.A.; Cross, G.A.M.
Mol. Biochem. Parasitol. 2, 135-150, 1981
A/Title: Glycoproteins from variant surface glycoproteins of *Trypanosoma brucei*. C-termir
A/Reference number: A61512; PMID:61172835; PMID:6163983
A/Accession: C61512
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-8 <HOL>
A/Cross-references: UNIPROT:Q7M3S3
C/Keywords: glycoprotein

Query Match 23.2%; Score 23; DB 2; Length 8;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 6 WNNN 9
Db 1 WENN 4

RESULT 15
A41117
acetylcholinesterase (EC 3.1.1.7), venom - Asian cobra (fragment)
C/Species: *Naja naja oxiana* (Asian cobra, Oxus cobra)
C/Date: 27-Mar-1992 #sequence_revision 27-Mar-1992 #text_change 09-Jul-2004
C/Accession: A41117
R/Kreienkamp, H.J.; Weise, C.; Raba, R.; Avilksaar, A.; Huch, F.
Proc. Natl. Acad. Sci. U.S.A. 88, 6117-6121, 1991
A/Title: Antonic subites of the catalytic center of acetylcholinesterase from Torpedo ar
A/Reference number: A41117; PMID:91296772; PMID:2068091
A/Accession: A41117
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-8 <KRB>
A/Cross-references: UNIPROT:Q7LZ27
C/Keywords: carboxylic ester hydrolase

Query Match 23.2%; Score 23; DB 2; Length 8;
Best Local Similarity 60.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 LMNNN 9

Db 4 MNPN 8

RESULT 16

IG heavy chain DJ region (clone C100-91A) - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
 C:Accession: PH1344
 R:Masseraman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
 J. Exp. Med. 176, 1577-1581, 1992
 A:Title: Predominance of fetal type DJH joining in young children with B precursor lymph
 A:Reference number: PH1302; MUID:93094761; PMID:1460419
 A:Accession: PH1344
 A:Molecule type: DNA
 A:Residues: 1-10 <WAS>
 C:Keywords: heterotetramer; immunoglobulin

Query Match 23.2% Score 23; DB 2; Length 10;
 Best Local Similarity 80.0%; Pred. No. 9.6e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 SDGLW 6
 Db 2 SDGLW 6

RESULT 17

proboscipedia - fruit fly (Drosophila pseudoobscura) (fragment)
 C:Species: Drosophila pseudoobscura
 C:Date: 31-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
 C:Accession: A4874
 R:Randazzo, F.M.; Cribbs, D.L.; Kaufman, T.C.
 Development 113, 257-271, 1991
 A:Title: Rescue and regulation of proboscipedia: a homeotic gene of the Antennapedia Com
 A:Reference number: A4874; MUID:92111389; PMID:1684932
 A:Accession: A4874
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-12 <RAN>
 A:Cross-references: UNIPROT:026429; GB:S77929; NID:G242124; PIDN:AA820845.1; PID:G242125
 A:Note: sequence extracted from NCBI backbone (NCBIN:77929, NCBIPI:77931)
 C:Keywords: sequence extracted from NCBI backbone

Query Match 23.2% Score 23; DB 2; Length 12;
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 NNNQ 10
 Db 9 NNNQ 12

RESULT 18

TS1735
 T-cell receptor beta-chain joining region - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 07-May-1995 #sequence_revision 01-Sep-1995 #text_change 05-Nov-1999
 C:Accession: S51735
 R:Dunovic-Bello, I.; Steinfel, A.; Ziegler, A.G.; Schendel, D.J.
 submitted to the EMBL Data Library, November 1993
 A:Reference number: S51732
 A:Accession: S51735
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-15 <DUR>
 A:Cross-references: EMBL:228344; NID:g607122; PIDN:CAA82198.1; PID:g607123

C:Keywords: T-cell receptor

Query Match 23.2% Score 23; DB 2; Length 15;
 Best Local Similarity 60.0%; Pred. No. 1.5e+03;
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 SSDGLNNNQ 10
 Db 4 SSDRLGNQPO 13

RESULT 19

anantin - Streptomyces coereulens
 C:Species: Streptomyces coereulens
 C:Date: 03-May-1994 #sequence_revision 05-Apr-1995 #text_change 09-Jul-2004
 C:Accession: A61211
 R:Wyes, D.F.; Lahm, H.W.; Manneberg, M.; Labhardt, A.M.
 J. Antibiot. 44, 172-180, 1991
 A:Title: Anantin -- a peptide antagonist of the atrial natriuretic factor (ANF). II. Det
 A:Reference number: A61211; MUID:91185186; PMID:1826288
 A:Accession: A61211
 A:Molecule type: protein
 A:Residues: 1-17 <WYS>
 A:Cross-references: UNIPROT:Q7M0J9
 A:Note: the isopeptide linked residue 8 is shown as Asn rather than Asp
 P:1-8/Cross-link: isopeptide amino end (Gly-Asn) #status experimental

Query Match 23.2% Score 23; DB 2; Length 17;
 Best Local Similarity 75.0%; Pred. No. 1.7e+03;
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 6 MNN 9
 Db 5 MNN 8

RESULT 20

CS6211
 progesterone receptor-related protein p23 - rabbit (fragment)
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 20-Jun-2000
 C:Accession: CS6211
 R:Johnson, J.L.; Belito, T.G.; Krco, C.J.; Toft, D.O.
 Mol. Cell. Biol. 14, 1956-1963, 1994
 A:Title: Characterization of a novel 23-kilodalton protein of inactive progesterone rece
 A:Reference number: A56211; MUID:94158868; PMID:814727
 A:Accession: CS6211
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-18 <OH>
 C:Superfamily: Caenorhabditis elegans hypothetical protein ZC395.10
 C:Keywords: steroid hormone receptor

Query Match 23.2% Score 23; DB 2; Length 18;
 Best Local Similarity 27.3%; Pred. No. 1.8e+03;
 Matches 3; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Oy 6 MNNQTOLE 16
 Db 8 WYDRDYVFIE 18

RESULT 21

AA4854
 L-2,4-diaminobutyrate decarboxylase (BC 4.1.1.-) - Acinetobacter calcoaceticus (fragment)
 N:Alternate names: DABA decarboxylase
 C:Species: Acinetobacter calcoaceticus
 C:Date: 31-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
 C:Accession: A44854; A41817
 R:Yamamoto, S.; Teuszaki, Y.; Tougou, K.; Shinoda, S.
 J. Gen. Microbiol. 138, 1461-1465, 1992
 A:Title: Purification and characterization of L-2,4-diaminobutyrate decarboxylase from A

A:Reference number: A44854; MUID:92381494; PMID:1512577
 A:Accession: A44854
 A:Molecule type: protein
 A:Residues: 1-19 <YAM>
 A:Cross-references: UNIPROT:O9P519
 A:Experimental source: ATCC 23055
 A>Note: sequence extracted from NCBI backbone (NCBIF:112311)
 C:Keywords: carbon-carbon lyase; carboxy-lyase

Query Match 23.2%; Score 23; DB 2; Length 19;
 Best Local Similarity 42.9%; Pred. No. 2e+03;
 Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 LMNNQQT 11
 |||:
 Db 11 LMNDAS 17

RESULT 22

S16202
 Pyruvate-5-carboxylate reductase - soybean chloroplast

C:Species: chloroplast Glycine max (soybean)

C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Aug-1997

C:Accession: S16202
 R:Chilson, O.P.; Kelly-Chilson, A.E.; Siegel, N.R.

Arch. Biochem. Biophys. 288, 350-357, 1991

A:Title: Pyruvate-5-carboxylate reductase in soybean nodules: isolation/partial primary
 A:Reference number: S16202; MUID:91378472; PMID:1898034

A:Accession: S16202

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-20 <ARC>

C:Genetics:

A:Genome: chloroplast

C:Keywords: chloroplast

Query Match 23.2%; Score 23; DB 2; Length 20;
 Best Local Similarity 30.0%; Pred. No. 2.1e+03;
 Matches 3; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 SDGLMNNQQT 11
 :|||:
 Db 7 AEGQMKDDVT 16

RESULT 23

S56756

link protein - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 17-Mar-1999

C:Accession: S56756

R:Rhodes, C.; Yamada, Y.

Nucleic Acids Res. 23, 2305-2313, 1995

A:Title: Characterization of a glucocorticoid responsive element and identification of a

A:Reference number: S56756; MUID:95334387; PMID:7610060

A:Accession: S56756

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-20 <RHO>

A:Cross-references: EMBL:X55057

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1991

Query Match 23.2%; Score 23; DB 2; Length 20;
 Best Local Similarity 60.0%; Pred. No. 2.1e+03;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 WNNNQ 10
 |||:
 Db 16 WNSGQ 20

RESULT 24

PH0923

T-cell receptor beta chain V-D-J region (isolate 9) - rat (fragment)
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 30-May-1997

C:Accession: PH0923

R:Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenberg, A.A.; Wilson, D.B.

J. Exp. Med. 174, 1467-1476, 1991

A:Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergic

A:Reference number: PH0891; MUID:92078857; PMID:1836012

A:Accession: PH0923

A:Molecule type: mRNA

A:Residues: 1-10 <GOL>

A:Experimental source: concanavalin A-activated lymphoblast

C:Keywords: T-cell receptor

Query Match 22.2%; Score 22; DB 2; Length 10;
 Best Local Similarity 60.0%; Pred. No. 1.4e+03;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 WNNNQ 10
 |||:
 Db 4 WGRNQ 8

RESULT 25

PH1471
 T-cell receptor beta chain (clone A24/PEF4) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 15-Mar-2004

C:Accession: PH1471

R:Caranova, J.L.; Martinon, F.; Gournier, H.; Barra, C.; Pannetier, C.; Regnault, A.; Koi

J. Exp. Med. 177, 811-820, 1993

A:Title: T cell receptor selection by and recognition of two class I major histocompatibility

A:Reference number: PH1430; MUID:93171821; PMID:8436911

A:Accession: PH1471

A:Molecule type: mRNA

A:Residues: 1-14 <CAS>

A:Experimental source: cytolytic T-lymphocyte

C:Keywords: receptor; T-cell

Query Match 22.2%; Score 22; DB 2; Length 14;
 Best Local Similarity 50.0%; Pred. No. 2e+03;
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 NNNQQLP 14
 :|||:
 Db 6 DNQDTQYF 13

RESULT 26

G49655
 T-cell-receptor beta chain variable region (clone 1) - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 30-May-1997

C:Accession: G49655; P49655; A49655

R:Gizom, A.A.; Thompson, S.D.; Luytink, L.; Passo, M.; Choi, E.; Glaas, D.N.

Proc. Natl. Acad. Sci. U.S.A. 90, 11104-11108, 1993

A:Title: Dominant T-cell-receptor beta chain variable region V beta 14+ clones in juvenile

A:Reference number: A49655; MUID:94068553; PMID:8248215

A:Accession: G49655

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 1-15 <GR1>

A:Experimental source: hip joint, synovial tissues

A>Note: this was designated clone 1

A>Note: sequence extracted from NCBI backbone (NCBIF:140453)

A:Accession: P49655

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 1-15 <GR2>

A:Experimental source: hip joint, synovial tissues

A>Note: sequence extracted from NCBI backbone (NCBIF:140451)

A>Note: this was designated clone 2

A:Accession: A49655

A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-15 <GRO>
A:Experimental source: knee joint, synovial fluid lymphocytes
A>Note: this was designated clone SF-1
A>Note: sequence extracted from NCBI backbone (NCBI:140445)
C:Keywords: T-cell receptor

Query Match 22.2% Score 22; DB 2; Length 15;
Best Local Similarity 40.0% Pred. No. 2.2e+03;
Matches 4; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
OY 1 SSDGLMNNQ 10
DB 5 SPDAGWTDQ 14

RESULT 27
B45474
chromopondin 2 - bovine (fragment)
N:Alternate names: corticotrophin-induced secreted protein (CISP); thrombospondin homolog
C:Species: Bos primigenius laurus (cattle)
C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: B45474
R:Peilert, S.; Lafenille, B.; Scherrer, N.; Gagnon, J.; Shi, D.L.; Chambaz, E.M.; Fei
J. Biol. Chem. 268, 4304-4310, 1993
A:Title: Corticotrophin-induced secreted protein, an ACTH-induced protein secreted by adr
A:Reference number: A45474; MUID:93179438; PMID:8382699
A:Accession: B45474
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-15 <PEL>
A:Cross-references: UNIPROT:Q9T597
A:Experimental source: adrenocortical cells
A>Note: sequence extracted from NCBI backbone (NCBI:125842)

Query Match 22.2% Score 22; DB 2; Length 15;
Best Local Similarity 80.0% Pred. No. 2.2e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 8 NNQTO 12
DB 2 NNQAO 6

RESULT 28
S57519
T cell receptor beta chain V region - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 10-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999
C:Accession: S57519
R:Burrows, S.R.; Silins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Argset, V.P.
submitted to the EMBL Data Library, June 1995
A:Description: T cell receptor repertoire for a viral epitope in humans is diversified b
A:Reference number: S57494
A:Accession: S57519
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-17 <BUR>
A:Cross-references: EMBL:249930; NID:9887494; PIDN:CAA90176.1; PID:9887495
C:Keywords: T-cell receptor

Query Match 22.2% Score 22; DB 2; Length 17;
Best Local Similarity 42.9% Pred. No. 2.5e+03;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
OY 1 SSDGLMNNQ 14
DB 3 SSGLSISSEYQYF 16

RESULT 29
S57556

T cell receptor beta chain V-D-J region (clone PP7 and clone TFL1) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 19-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999
C:Accession: S57556; S57557
R:Burrows, S.R.; Silins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Argset, V.P.
submitted to the EMBL Data Library, June 1995
A:Description: T cell receptor repertoire for a viral epitope in humans is diversified b
A:Reference number: S57494
A:Accession: S57556
A:Molecule type: mRNA
A:Residues: 1-17 <BUR>
A:Cross-references: EMBL:249926; NID:9887498; PIDN:CAA90172.1; PID:9887499
A:Experimental source: clone PP7
A:Accession: S57557
A:Molecule type: mRNA
A:Residues: 1-17 <BUR>
A:Cross-references: EMBL:249928; NID:9887502; PIDN:CAA90174.1; PID:9887503
A:Experimental source: clone TFL1
C:Keywords: T-cell receptor

Query Match 22.2% Score 22; DB 2; Length 17;
Best Local Similarity 42.9% Pred. No. 2.5e+03;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
OY 1 SSDGLMNNQ 14
DB 3 SSGLSISSEYQYF 16

RESULT 30
I40062
shikimate 5-dehydrogenase (EC 1.1.1.25) - Buchnera aphidicola (fragment)
C:Species: Buchnera aphidicola
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 08-Oct-1999
C:Accession: I40062
R:Roubakshin, D.; Baumann, P.
Gene 155, 107-112, 1995
A:Title: Characterization of a putative 23S-5S rRNA operon of Buchnera aphidicola (endos
A:Reference number: I40061; MUID:95212914; PMID:7535281
A:Accession: I40062
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-18 <RES>
A:Cross-references: EMBL:U10496; NID:9854711; PIDN:AAA9125.1; PID:9854712
C:Genetics:
A:Gene: aroE
C:Keywords: oxidoreductase

Query Match 22.2% Score 22; DB 2; Length 18;
Best Local Similarity 75.0% Pred. No. 2.7e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 5 LMNN 8
DB 4 LMHN 7

RESULT 31
PC1251
testin II - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Feb-1997
C:Accession: PC1251
R:Cheng, C.Y.; Morris, I.; Bardin, C.W.
Biochem. Biophys. Res. Commun. 191, 224-231, 1993
A:Title: Testins are structurally related to the mouse cysteine proteinase precursor but
A:Reference number: PC1250; MUID:93191698; PMID:8447824
A:Accession: PC1251
A:Molecule type: protein
A:Residues: 1-19 <CHR>
C:Superfamily: papain

Query Match 22.2% Score 22; DB 2; Length 19;

Best Local Similarity 41.7%; Pred. No. 2.8e+03;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 1 SSDGLMNNQTO 12
Db 8 SLDVEMNEMRKT 19

RESULT 32

140063

shikimate 5-dehydrogenase (EC 1.1.1.25) - Buchnera aphidicola (fragment)

C/Species: Buchnera aphidicola
C/Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 09-Jul-2004

C/Accession: 140063

R/Rouhakhsh, D.; Baumann, P.

Gene 155, 107-112, 1995

A/Title: Characterization of a putative 23S-5S rRNA operon of Buchnera aphidicola (endosymbiont)

A/Reference number: 140061; MUID:95212914; PMID:7535281

A/Accession: 140063

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-19 <RES>

A/Cross-references: UNIPROT:Q4608; EMBL:U10497; NID:g854713; PIDN:AAA79126.1; PID:g8547

C/Genetics: A;Gene: aroB

C/Keywords: oxidoreductase

Query Match 22.2%; Score 22; DB 2; Length 19;
Best Local Similarity 75.0%; Pred. No. 2.8e+03;

Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 LMNN 8

Db 4 LMND 7

RESULT 33

H22565

R-phycocerythrin gamma-C chain - red alga (Gaetoclonium coulteri) (fragment)

C/Species: Gaetoclonium coulteri

C/Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 09-Jul-2004

C/Accession: H22565

R/Klotz, A.V.; Glazer, A.N.

J. Biol. Chem. 260, 4856-4863, 1985

A/Title: Characterization of the bilin attachment sites in R-phycocerythrin.

A/Reference number: A22565; MUID:85182601; PMID:3886644

A/Accession: H22565

A/Molecule type: protein

A/Residues: 1-20 <KLO>

A/Cross-references: UNIPROT:Q7M270

C/Superfamily: Aglaothamnion neglectum R-phycocerythrin gamma chain 33

Query Match 22.2%; Score 22; DB 2; Length 20;
Best Local Similarity 37.5%; Pred. No. 3e+03;

Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 7 NNNQTOLE 14

Db 8 HHEEQIF 15

RESULT 34

A39308

glycine reductase (EC 1.4.99.-) sulphydryl protein C, alpha chain - Clostridium sticklandii

C/Species: Clostridium sticklandii

C/Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 09-Jul-2004

C/Accession: A39308

R/Stadman, T.C.; Davis, J.N.

J. Biol. Chem. 266, 22147-22153, 1991

A/Title: Glycine reductase protein C. Properties and characterization of its role in the

A/Reference number: A39308; MUID:92042141; PMID:1935235

A/Accession: A39308

A/Status: preliminary

A/Molecule type: protein
A/Residues: 1-8 <STA>
A/Cross-references: UNIPROT:Q7M0L0
C/Function:
A/Description: glycine reductase complex catalyzes the reductive deamination of glycine to
C/Keywords: ATP; oxidoreductase

Query Match 21.2%; Score 21; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LMN 7

Db 5 LMN 7

RESULT 35

PT0562

T-cell receptor beta chain V-D-J region (126-1AK) - mouse (fragment)

C/Species: Mus musculus (house mouse)

C/Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C/Accession: PT0562

R/Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A/Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A/Reference number: PT0509; MUID:9127601; PMID:1711558

A/Accession: PT0562

A/Status: translation not shown

A/Residues: 1-9 <FE>

A/Experimental source: day 18 fetal thymus, strain BALB/c

C/Keywords: T-cell receptor

Query Match 21.2%; Score 21; DB 2; Length 9;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSDGLW 6

Db 2 SSDNMW 7

RESULT 36

PH1308

Ig heavy chain DJ region (clone C731-94) - human (fragment)

C/Species: Homo sapiens (man)

C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C/Accession: PH1308

R/Wasserman, R.; Gallili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.

J. Exp. Med. 176, 1577-1581, 1992

A/Title: Predominance of fetal type DJH joining in young children with B precursor lymphoma

A/Reference number: PH1302; MUID:93094761; PMID:1460419

A/Accession: PH1308

A/Molecule type: DNA

A/Residues: 1-12 <WAS>

C/Keywords: heterotetramer; immunoglobulin

Query Match 21.2%; Score 21; DB 2; Length 12;
Best Local Similarity 75.0%; Pred. No. 2.5e+03;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GLMN 7

Db 8 GQNM 11

RESULT 37

S57572

T cell receptor V-J junctional alpha chain region - human (fragment)

C/Species: Homo sapiens (man)

C/Date: 19-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999

C/Accession: S57572

R/Burrows, S.R.; Silins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Arguet, V.P.

submitted to the EMBL Data Library, June 1995

A:Description: T cell receptor repertoire for a viral epitope in humans is diversified

A:Reference number: S57494

A:Accession: S57572

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-14 <BUR>

A:Cross-references: EMBL:249957; NID:g887478; PIDN:CAA90228.1; PID:g887479

C:Keywords: T-cell receptor

Query Match 21.2%; Score 21; DB 2; Length 14;

Best Local Similarity 33.3%; Pred. No. 2.9e+03;

Matches 3; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 6 WNNNOTOLF 14

DB 5 WSGNTPLVF 13

RESULT 38

PH1594

Ig H chain V-D-J region (wild-type clone 149) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999

C:Accession: PH1594

R:Levinson, D.A.; Campos-Torres, J.; Leder, P.

J. Exp. Med. 178, 317-329, 1993

A:Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice

A:Reference number: PH1580; MUID:93101609; PMID:8315387

A:Accession: PH1594

A:Molecule type: DNA

A:Residues: 1-14 <LEV>

A:Experimental source: bone marrow pre-B lymphocyte

C:Keywords: immunoglobulin

Query Match 21.2%; Score 21; DB 2; Length 14;

Best Local Similarity 100.0%; Pred. No. 2.9e+03;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GLW 6

DB 11 GLW 13

RESULT 39

PH1594

alpha-conotoxin MII - cone shell (Conus magus)

C:Species: Conus magus (magus cone)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004

C:Accession: A59046

R:Cartier, G.E.; Yoshikami, D.; Gray, W.R.; Luo, S.; Olivera, B.M.; McIntosh, J.M.

J. Biol. Chem. 271, 7522-7528, 1996

A:Title: A new alpha-conotoxin which targets alpha3beta2 nicotinic acetylcholine receptor

A:Reference number: A59046; MUID:96205934; PMID:8631783

A:Accession: A59046

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-16 <CAR>

A:Cross-references: UNIPROT:P56636

C:Superfamily: alpha-conotoxin

C:Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; postsynaptic neuro

F:1-16/Product: alpha-conotoxin MII #status experimental <MAT>

F:16/Modified site: amidated carboxyl end (Cys) #status experimental

Query Match 21.2%; Score 21; DB 2; Length 16;

Best Local Similarity 100.0%; Pred. No. 3.4e+03;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 LEHS 18

DB 10 LEHS 13

RESULT 40

S28213

glutathione transferase (EC 2.5.1.18) - European toad (fragments)

C:Species: Bufo bufo (European toad)

C:Date: 19-Mar-1997 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999

C:Accession: S28213

R:Accero, A.; Dragani, B.; Bucciarrelli, T.; Sacchetta, P.; Martini, F.; Angelucci, S.; Am

Biochem. J. 289, 417-422, 1993

A:Title: Purification and characterization of the major glutathione transferase from adu

A:Reference number: S28213; MUID:93143709; PMID:8424786

A:Accession: S28213

A:Molecule type: protein

A:Residues: 1-16 <ACE>

A:Experimental source: liver

C:Keywords: transferase

Query Match 21.2%; Score 21; DB 2; Length 16;

Best Local Similarity 45.5%; Pred. No. 3.4e+03;

Matches 5; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

OY 6 WNNNOTOLF 16

DB 4 WDNEAN--FLB 12

Search completed: December 30, 2004, 16:32:07

Job time : 40 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 16:14:44 ; Search time 188 Seconds
(without alignments)
55.089 Million cell updates/sec

Title: US-10-718-321-1

Perfect score: 99

Sequence: 1 SSDGLMNNNTQLFLHS 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 14317

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : UniProt 02:*

1: uniprot_sprotc:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	30.3	19	2	Q62545-mus spreus
2	27	27.3	20	2	Q65705-bovine resp
3	27	27.3	20	2	Q8V686-bovine resp
4	26	26.3	17	2	Q78378-human immun
5	26	26.3	19	2	Q9ETX8-helicobacte
6	26	26.3	20	2	Q6SBY5-drosophila
7	26	26.3	20	2	AA884726-helicobacte
8	25	25.3	10	2	Q9TR47-bos taurus
9	25	25.3	19	2	Q15610-homo sapien
10	25	25.3	19	2	Q9ETX7-helicobacte
11	25	25.3	19	2	Q84197-human resp
12	25	25.3	19	2	Q91329-human immun
13	25	25.3	19	2	Q9WJBI-human immun
14	25	25.3	20	1	A1BG-EQUUS
15	25	25.3	20	2	Q9URC7-baccharomy
16	24.5	24.7	20	2	Q9R987-nitrosospi
17	24.5	24.7	20	2	Q9R985-nitrosospi
18	24	24.2	11	2	Q6YBE8-streptococ
19	24	24.2	11	2	AA060518-streptoc
20	24	24.2	11	2	AA060520-streptoc
21	24	24.2	11	2	AA060522-streptoc
22	24	24.2	11	2	AA060524-streptoc
23	24	24.2	11	2	AA060526-streptoc
24	24	24.2	11	2	AA060528-streptoc
25	24	24.2	11	2	AA060530-streptoc
26	24	24.2	13	2	Q7TMB4-mus musculi
27	24	24.2	16	2	Q9R557-bacillus sp
28	24	24.2	16	2	Q66737-equine infe
29	24	24.2	17	2	Q81VK5-homo sapien
30	24	24.2	17	2	Q78381-human immun
31	24	24.2	18	2	Q9HBD6-homo sapien
32	24	24.2	18	2	Q9ZG65-chlamydia t
33	24	24.2	19	2	Q9QUY4-rattus sp.
34	24	24.2	19	2	Q9QVJ2-rattus sp.
35	24	24.2	20	1	LRH3_HUMAN
36	24	24.2	20	1	OMP_HAEGA
37	24	24.2	20	1	OMP_HAEGA
38	24	24.2	20	2	OS1950-chlamydia t
39	24	24.2	20	2	OS1952-chlamydia t
40	24	24.2	20	2	Q7TMA1-mus musculi
41	24	24.2	20	2	Q90X92-gallus gall
42	23.5	23.7	19	1	DCAM_ACACA
43	23.5	23.7	20	2	Q9UC28-homo sapien
44	23	23.2	8	2	Q7M363-trypanosoma
45	23	23.2	8	2	Q7LZ27-naja oxiata
46	23	23.2	11	2	Q6LBJ0-mus musculi
47	23	23.2	11	2	CA50294-mus musculi
48	23	23.2	12	2	Q64429-drosophila
49	23	23.2	14	2	Q84072-influenza a
50	23	23.2	15	1	DC16_MALZE
51	23	23.2	16	2	Q9S8D3-cynara card
52	23	23.2	16	2	Q75710-human immun
53	23	23.2	17	1	ANAN_STRCD
54	23	23.2	18	2	P82674-bos taurus
55	23	23.2	18	2	Q9TRB6-bos taurus
56	23	23.2	18	2	Q94PS2-macropodus
57	23	23.2	18	2	Q73920-human immun
58	23	23.2	18	2	Q73921-human immun
59	23	23.2	19	2	Q9R519-acinetobact
60	23	23.2	19	2	Q935K9-salmonella
61	23	23.2	20	1	DC1_CANFA
62	23	23.2	20	2	Q7BVQ6-canis, fami
63	23	23.2	20	2	Q9R4U9-mycoplasma
64	23	23.2	20	2	Q73917-acinetobact
65	23	23.2	20	2	Q73918-human immun
66	23	23.2	20	2	Q78486-human immun
67	23	23.2	20	2	Q78505-human immun
68	22	22.2	8	2	Q81VK3-homo sapien
69	22	22.2	8	2	Q85406-coxiella bu
70	22	22.2	12	2	Q9GE05-etabailia g
71	22	22.2	12	2	Q6LEB4-plasmid rk2
72	22	22.2	12	2	AA15046-plasmid r
73	22	22.2	15	2	Q9QUY7-mus sp. . g
74	22	22.2	16	2	Q7W227-pseudomonas
75	22	22.2	17	2	Q6LC11-mus musculi
76	22	22.2	17	2	Q78327-human immun
77	22	22.2	17	2	Q78380-human immun
78	22	22.2	17	2	AA6C3052-mus muscu
79	22	22.2	18	2	Q96RV9-homo sapien
80	22	22.2	18	2	Q6LDD4-rattus sp.
81	22	22.2	18	2	Q9PX66-homo adeno
82	22	22.2	18	2	AA837716-rattus sp
83	22	22.2	19	2	Q96CE2-homo sapien
84	22	22.2	19	2	Q44608-buchnera ap
85	22	22.2	20	2	Q7PEW6-anopheles g
86	22	22.2	20	2	Q7M270-gastrocloni
87	22	22.2	20	2	Q9QUY9-rattus sp.
88	21	21.2	7	2	Q9BRV4-homo sapien
89	21	21.2	8	2	Q7W0U0-clostridium
90	21	21.2	8	2	AA833374-homo sapi
91	21	21.2	9	2	Q6LDD7-staphylococ
92	21	21.2	9	2	AA26674-staphyloc
93	21	21.2	12	2	Q8H6E6-hordeum vul
94	21	21.2	14	1	UN46_CLOA
95	21	21.2	14	2	Q6Q6T3-strephanandr
96	21	21.2	14	2	Q6Q6Y5-strephanandr
97	21	21.2	14	2	Q6Q6V7-strephanandr
98	21	21.2	14	2	Q6Q6V9-neillia uek
99	21	21.2	14	2	Q6Q6Z2-neillia chi
100	21	21.2	14	2	Q6Q6Z4-neillia chi
101	21	21.2	14	2	Q6Q6Z5-neillia spa
102	21	21.2	14	2	Q6Q6Z7-neillia sin
103	21	21.2	14	2	Q6Q6Z8-neillia af
104	21	21.2	14	2	Q6Q702-physocarpus

```

105      21      21.2      14      2      06Q704      06Q704 physocarpus
106      21      21.2      14      2      06Q706      06Q706 physocarpus
107      21      21.2      14      2      06Q709      06Q709 physocarpus
108      21      21.2      14      2      06Q710      06Q710 physocarpus
109      21      21.2      14      2      06Q712      06Q712 physocarpus
110      21      21.2      14      2      06Q713      06Q713 vaquellinia
111      21      21.2      14      2      06Q714      06Q714 lyonothamu
112      21      21.2      14      2      AAS65671      AAS65671 lyonothamu
113      21      21.2      14      2      AAS65672      AAS65672 vaquellinia
114      21      21.2      14      2      AAS65673      AAS65673 vaquellinia
115      21      21.2      14      2      AAS65674      AAS65674 physocarp
116      21      21.2      14      2      AAS65675      AAS65675 physocarp
117      21      21.2      14      2      AAS65676      AAS65676 physocarp
118      21      21.2      14      2      AAS65677      AAS65677 physocarp
119      21      21.2      14      2      AAS65678      AAS65678 physocarp
120      21      21.2      14      2      AAS65679      AAS65679 physocarp
121      21      21.2      14      2      AAS65680      AAS65680 physocarp
122      21      21.2      14      2      AAS65681      AAS65681 physocarp
123      21      21.2      14      2      AAS65682      AAS65682 physocarp
124      21      21.2      14      2      AAS65683      AAS65683 physocarp
125      21      21.2      14      2      AAS65684      AAS65684 physocarp
126      21      21.2      14      2      AAS65685      AAS65685 physocarp
127      21      21.2      14      2      AAS65686      AAS65686 physocarp
128      21      21.2      14      2      AAS65687      AAS65687 neillia a
129      21      21.2      14      2      AAS65688      AAS65688 neillia s
130      21      21.2      14      2      AAS65689      AAS65689 neillia s
131      21      21.2      14      2      AAS65690      AAS65690 neillia c
132      21      21.2      14      2      AAS65691      AAS65691 neillia c
133      21      21.2      14      2      AAS65692      AAS65692 neillia c
134      21      21.2      14      2      AAS65693      AAS65693 neillia c
135      21      21.2      14      2      AAS65694      AAS65694 neillia c
136      21      21.2      14      2      AAS65695      AAS65695 neillia c
137      21      21.2      14      2      AAS65696      AAS65696 neillia u
138      21      21.2      14      2      AAS65697      AAS65697 neillia u
139      21      21.2      14      2      AAS65698      AAS65698 stephanan
140      21      21.2      14      2      AAS65699      AAS65699 stephanan
141      21      21.2      14      2      AAS65700      AAS65700 stephanan
142      21      21.2      14      2      AAS65701      AAS65701 stephanan
143      21      21.2      14      2      AAS65702      AAS65702 stephanan
144      21      21.2      14      2      AAS65703      AAS65703 stephanan
145      21      21.2      15      2      Q7S007      Q7S007 neurospora
146      21      21.2      15      2      O6LAG2      O6LAG2 oryctolagus
147      21      21.2      15      2      CAAT7016      CAAT7016 oryctolagus
148      21      21.2      17      2      062436      062436 mus musculu
149      21      21.2      17      2      O9EP50      O9EP50 unidentified
150      21      21.2      18      1      TOP1_KLEAE      P46155 klebsiella

```

ALIGNMENTS

```

RESULT 1
ID      062545      PRELIMINARY;      PRT;      19 AA.
AC      062545;
DT      01-JUN-1998 (TReMBLrel. 06, Created)
DT      01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT      01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE      Lactate dehydrogenase-C (Fragment).
OS      Mus spretus (Western wild mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10096;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=SPRET/E1;
RX      MEDLINE=9431082; PubMed=8043349;
RA      Ko M.S., Wang X., Horton J.H., Hagen M.D., Takahashi N., Maezaki Y.,
RA      Nadeau J.H.;
RT      "Genetic mapping of 40 cDNA clones on the mouse genome by PCR.";
RL      Mamm. Genome 5:349-355(1994).
DR      EMBL; U05739; AAB60478.1; -.
DR      PIR; I49422; I49422.

```

```

DR      HSSP; P00342; 2LUX.
FT      NON TER      1
SQ      SEQUENCE      19 AA; 2203 MW; 80DB2B6F765516B CRC64;
Query Match      30.3%; Score 30; DB 2; Length 19;
Best Local Similarity 71.4%; Pred. No. 9.9e+02;
Matches      5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy      1 SSDGLMN 7
      | : | | |
Db      6 SADTLMN 12

```

RESULT 2

```

ID      065705      PRELIMINARY;      PRT;      20 AA.
AC      065705;
DT      01-NOV-1996 (TReMBLrel. 01, Created)
DT      01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT      01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE      RNA-dependent RNA polymerase (Fragment).
GN      Name-L protein;
OS      Bovine respiratory syncytial virus.
OC      Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC      Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX      NCBI_TaxID=11246;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Zamora M., Samal S.K.;
RL      Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.
DR      EMBL; M82816; AAA42807.1; -.
DR      PIR; P00300; P00300.
DR      GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
KW      RNA-directed RNA polymerase.
FT      NON TER      20
SQ      SEQUENCE      20 AA; 2358 MW; 4F13BA084822ED09 CRC64;

```

```

Query Match      27.3%; Score 27; DB 2; Length 20;
Best Local Similarity 37.5%; Pred. No. 3.1e+03;
Matches      6; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
Qy      3 DGLMNNQTLFLEHS 18
      | : | | | : |
Db      2 DTLIHENSTWVYLTDS 17

```

RESULT 3

```

ID      Q8V686      PRELIMINARY;      PRT;      20 AA.
AC      Q8V686;
DT      01-MAR-2002 (TReMBLrel. 20, Created)
DT      01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT      01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE      RNA-dependent RNA polymerase major subunit (Fragment).
OS      Bovine respiratory syncytial virus (strain A51908) (BRV).
OC      Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC      Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX      NCBI_TaxID=11247;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=A51908;
RX      MEDLINE=21580793; PubMed=11724268;
RA      Yunus A.S., Khatrar S.K., Collins P.L., Samal S.K.;
RT      "Rescue of bovine respiratory syncytial virus from cloned cDNA: entire
RT      genome sequence of BRV strain A51908.";
RL      Virus Genes 23:157-164(2001).
DR      EMBL; AF295544; AAL49413.1; -.
DR      PIR; P00300; P00300.
DR      GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
KW      RNA-directed RNA polymerase.
FT      NON TER      20
SQ      SEQUENCE      20 AA; 2358 MW; 4F13BA084822ED09 CRC64;

```

Query Match 27.3%; Score 27; DB 2; Length 20;
 Best Local Similarity 37.5%; Pred. No. 3.1e+03;
 Matches 6; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 3 DGLMNNQTLFLSHS 18
 Db 2 DTLIHENSTNYLTDS 17

RESULT 4

Q78378 PRELIMINARY; PRT; 17 AA.

AC Q78378;
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-1996 (TREMBLrel. 26, Last annotation update)
 DE Immunodeficiency virus type 1, viral sample FLBRS5A (Florida patient B), partial env cds, V5 region. (Fragment).
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.
 NCBI_TaxID=11676;
 RX MEDLINE=92271245; PubMed=1589796;
 RA Ou C.-Y., Ciesielski C.A., Myers G., Banda C.I., Luo C.C.,
 RA Korber B.T.M., Mullins J.I., Schochetman G., Berkman R.L.,
 RA Economou A.N., Witte J.J., Furman L.J., Satten G.A., Curran J.W.,
 RA Jaffe H.W.;
 RL "Molecular epidemiology of HIV transmission in a dental practice";
 RT Science 256:1165-1171(1992).
 RN [2]

RP SEQUENCE FROM N.A.

RA Zhang L.Q., Leign-Brown A.J.;
 RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
 DR EMBL; M92123; AAA44493.1;
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR InterPro; IPR000777; GP120.
 FT NON TER 1 1
 FT SEQUENCE 17 AA; 1723 MW; 34757935D12CA370 CRC64;

Query Match 26.3%; Score 26; DB 2; Length 17;
 Best Local Similarity 50.0%; Pred. No. 3.8e+03;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 7 NNNQTLF 14
 Db 4 NTNETETP 11

RESULT 5

Q9ETX8 PRELIMINARY; PRT; 19 AA.

AC Q9ETX8;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE Glutamate racemase (Fragment).
 GN Name=glr;
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
 OC Helicobacteraceae; Helicobacter.
 NCBI_TaxID=210;
 RX MEDLINE=2117015; PubMed=1119371;
 RA Occhiali A., Marais A., Urdaci M., Sierra R., Munoz N., Covacci A.,
 RA Megraud F.;
 RL "Composition and gene expression of the cag pathogenicity island in RT Helicobacter pylori strains isolated from gastric carcinoma and RT gastritis patients in Costa Rica";
 RL Infect. Immun. 69:1902-1908(2001).

DR EMBL; AF289397; AAC09844.1; -
 DR EMBL; AF289390; AAC09830.1; -
 DR EMBL; AF289391; AAC09832.1; -
 DR EMBL; AF289393; AAC09836.1; -
 DR EMBL; AF289394; AAC09838.1; -
 DR EMBL; AF289395; AAC09840.1; -
 DR EMBL; AF289396; AAC09842.1; -
 FT NON TER 1 1
 FT SEQUENCE 19 AA; 2252 MW; 6FF0977B6A7CE7C5 CRC64;

Query Match 26.3%; Score 26; DB 2; Length 19;
 Best Local Similarity 31.2%; Pred. No. 4.3e+03;
 Matches 5; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 SSDGLMNNQTLFLE 16
 Db 3 SGDIVMLKQAKWK 18

RESULT 6

Q6SBY5 PRELIMINARY; PRT; 20 AA.

AC Q6SBY5;
 DT 05-JUN-2004 (TREMBLrel. 27, Created)
 DT 05-JUN-2004 (TREMBLrel. 27, Last sequence update)
 DT 05-JUN-2004 (TREMBLrel. 27, Last annotation update)
 DE Epidermal growth factor receptor (Fragment).
 GN Name=Egfr; ORFNames=CG1079;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RX MEDLINE=7227;
 RA [1]

RP SEQUENCE FROM N.A.

RA Palsson A., Rouse A., Riley-Berger R., Dworkin I., Gibson G.;
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY460749; AAR84726.1;
 DR GO; GO:0004872; F:receptor activity; IEA.
 FT NON TER 20 20
 FT SEQUENCE 20 AA; 2332 MW; E3AC72F1B8FE357 CRC64;

Query Match 26.3%; Score 26; DB 2; Length 20;
 Best Local Similarity 50.0%; Pred. No. 4.5e+03;
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SDGLMNNN 9
 Db 11 SRGLMDS 18

RESULT 7

AAR84726 PRELIMINARY; PRT; 20 AA.

AC AAR84726;
 DT 02-MAR-2004 (TREMBLrel. 27, Created)
 DT 02-MAR-2004 (TREMBLrel. 27, Last sequence update)
 DT 02-MAR-2004 (TREMBLrel. 27, Last annotation update)
 DE Epidermal growth factor receptor (Fragment).
 GN EGFR OR CG10079.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RX MEDLINE=7227;
 RA [1]

RP SEQUENCE FROM N.A.

RA Palsson A., Rouse A., Riley-Berger R., Dworkin I., Gibson G.;
 RL "Nucleotide variation in the EGFR locus of Drosophila melanogaster";
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY460749; AAR84726.1; -.
KM Recepor.
FT NON TER 20
SQ SEQUENCE 20 AA; 2332 MW; E3AC72F1BFP9E357 CRC64;

Query Match 26.3%; Score 26; DB 2; Length 20;
Best Local Similarity 50.0%; Pred. No. 4.5e+03;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 SDGLMNN 9
| | | | |
Db 11 SRGLWDS 18

RESULT 8

ID Q9TR47 PRELIMINARY; PRT; 10 AA.
AC Q9TR47,
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Amphoterin homolog (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=96029671; PubMed=7592757;
RA Hoti O., Brett J., Slatery T., Cao R., Zhang J., Chen J.X.,
RA Nagashima M., Lundh E.R., Vijay S., Nitecki D.,
RT "The receptor for advanced glycation end products (RAGE) is a cellular
binding site for amphoterin. Mediation of neurite outgrowth and co-
RT expression of rage and amphoterin in the developing nervous system.";
RL J. Biol. Chem. 270:25752-25761(1995).
SQ SEQUENCE 10 AA; 1163 MW; 28E5034453769B18 CRC64;

Query Match 25.3%; Score 25; DB 2; Length 10;
Best Local Similarity 75.0%; Pred. No. 3e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 LMNN 8
| | | | |
Db 5 MMNN 8

RESULT 9

ID Q15610 PRELIMINARY; PRT; 19 AA.
AC Q15610,
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Topoisomerase I (Fragment).
GN Name=TOP1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9109302; PubMed=2176592;
RA Kunze N., Klein M., Richter A., Koipipers R.,
RT "Structural characterization of the human DNA topoisomerase I gene
RT promoter";
RL Eur. J. Biochem. 194:323-330(1990).
DR EMBL; X52601; CA36834.1; -.
DR GO; GO:0016853; F:isomerase activity; IEA.
KM isomerase.
FT NON TER 19
SQ SEQUENCE 19 AA; 2236 MW; 47BDB36F44FFEFB CRC64;

Query Match 25.3%; Score 25; DB 2; Length 19;
Best Local Similarity 42.9%; Pred. No. 6.1e+03;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 SSDGLMNNOTQLE 14
| | | | |
Db 2 SGDHLHDSQLFF 15

RESULT 10

ID Q9ETX7 PRELIMINARY; PRT; 19 AA.
AC Q9ETX7,
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Glutamate racemase (Fragment).
GN Name=glr;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CR41, and CR44;
RX MEDLINE=21117015; PubMed=11179371;
RA Occhialini A., Marais A., Urdaci M., Sierra R., Munoz N., Covacci A.,
RA Megraud F.,
RT "Composition and gene expression of the cag pathogenicity island in
RT Helicobacter pylori strains isolated from gastric carcinoma and
RT gastritis patients in Costa Rica";
RL Infect. Immun. 69:1902-1908(2001).
DR EMBL; AF289398; AACG9846.1; -.
DR EMBL; AF289392; AACG9834.1; -.
FT NON TER 1
SQ SEQUENCE 19 AA; 2238 MW; 6FF0977B72DEF7C5 CRC64;

Query Match 25.3%; Score 25; DB 2; Length 19;
Best Local Similarity 31.2%; Pred. No. 6.1e+03;
Matches 5; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 SSDGLMNNOTQLE 16
| | | | |
Db 3 SGDVLMLKQKEMWLK 18

RESULT 11

ID Q84197 PRELIMINARY; PRT; 19 AA.
AC Q84197,
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JAN-1999 (TRENBLrel. 09, Last annotation update)
DE L protein (Fragment).
OS Human respiratory syncytial virus (subgroup B / strain 18537).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus;
OC Human respiratory syncytial virus B.
OX NCBI_TaxID=11251;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89036169; PubMed=3183631;
RA Johnson P.R., Collins P.L.,
RT "The A and B subgroups of human respiratory syncytial virus:
RT Comparison of intergenic and gene-overlap sequences";
RL J. Gen. Virol. 69:2901-2906(1988).
DR EMBL; D00397; BAA20958.1; -.
FT NON TER 19
SQ SEQUENCE 19 AA; 2100 MW; 8277C94820609735 CRC64;

Query Match 25.3%; Score 25; DB 2; Length 19;
Best Local Similarity 31.2%; Pred. No. 6.1e+03;
Matches 5; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

GO: GO:0005853; C:eukaryotic translation elongation factor 1 . . . ; IEA

DR GO:0006414; P:translational elongation; IEA.
 DR InterPro; IPR001326; EFL_BD.
 DR Pfam; PF00736; EFL_GNE; 1.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 20 AA; 2388 MW; 594377CBCE72B0D CRC64;
 Query Match 25.3%; Score 25; DB 2; Length 20;
 Best Local Similarity 37.5%; Pred. No. 6.5e+03;
 Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 6 MNNOQTOL 13
 DB 5 WDDDETNL 12

RESULT 16
 ID Q9R987 PRELIMINARY; PRT; 20 AA.
 AC Q9R987;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Ammonia monooxygenase subunit C1 (Fragment).
 GN Name=amoc1;
 OS Nitrosospirita sp. Np39-19.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
 OC Nitrosomonadaceae; Nitrosospirita.
 NCBI_TaxID=61907;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Np39-19;
 RX MEDLINE=21665709; PubMed=11807563;
 RA Norton J.M., Alzerreca J.J., Suwa Y., Klotz M.G.;
 RT "Diversity of ammonia monooxygenase operon in autotrophic ammonia-
 oxidizing bacteria." Arch. Microbiol. 177:139-149(2002).
 RL Arch. Microbiol. 177:139-149(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Np39-19;
 RX Shiozawa T.L., Norton J.M., Alzerreca J.J., Klotz M.G.;
 RL Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.
 DE EMBL; AF042170; AAC25054.1; -.
 DR GO:0004497; F:monooxygenase activity; IEA.
 KM Monooxygenase.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 20 AA; 2266 MW; A745DC66ACC3F2A9 CRC64;
 Query Match 24.7%; Score 24.5; DB 2; Length 20;
 Best Local Similarity 31.2%; Pred. No. 7.8e+03;
 Matches 5; Conservative 6; Mismatches 4; Indels 1; Gaps 1;
 QY 2 SDGLMNNNOQTOLPFLH 17
 DB 2 TDVVM-NGOSKVILNN 16

RESULT 17
 ID Q9R9A5 PRELIMINARY; PRT; 20 AA.
 AC Q9R9A5;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Ammonia monooxygenase 1 subunit C (Fragment).
 GN Name=amoc1;
 OS Nitrosospirita sp. NpAV.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
 OC Nitrosomonadaceae; Nitrosospirita.
 NCBI_TaxID=58133;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NpAV;

RX MEDLINE=96001263; PubMed=7557469;
 RA Klotz M.G., Norton J.M.;
 RT "Sequence of an ammonia monooxygenase subunit A-encoding gene from
 Nitrosospirita sp. NpAV." Gene 163:159-160(1995).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NpAV;
 RX MEDLINE=96275898; PubMed=8674986;
 RA Norton J.M., Low J.M., Klotz M.G.;
 RT "The gene encoding ammonia monooxygenase subunit A exists in three
 nearly identical copies in Nitrosospirita sp. NpAV." FEMS Microbiol. Lett. 139:181-188(1996).
 RL FEMS Microbiol. Lett. 139:181-188(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NpAV;
 RX MEDLINE=97306641; PubMed=9163908;
 RA Klotz M.G., Alzerreca J., Norton J.M.;
 RT "A gene encoding a membrane protein exists upstream of the amoA/amoB
 genes in ammonia oxidizing bacteria: a third member of the amo
 operon?" FEMS Microbiol. Lett. 150:65-73(1997).
 RL [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NpAV;
 RX MEDLINE=21665709; PubMed=11807563;
 RA Norton J.M., Alzerreca J.J., Suwa Y., Klotz M.G.;
 RT "Diversity of ammonia monooxygenase operon in autotrophic ammonia-
 oxidizing bacteria." Arch. Microbiol. 177:139-149(2002).
 RL Arch. Microbiol. 177:139-149(2002).
 DR EMBL; AF032458; AAB86880.1; -.
 DR GO:0004497; F:monooxygenase activity; IEA.
 KM Monooxygenase.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 20 AA; 2266 MW; A745DC66ACC3F2A9 CRC64;
 Query Match 24.7%; Score 24.5; DB 2; Length 20;
 Best Local Similarity 31.2%; Pred. No. 7.8e+03;
 Matches 5; Conservative 6; Mismatches 4; Indels 1; Gaps 1;
 QY 2 SDGLMNNNOQTOLPFLH 17
 DB 2 TDVVM-NGOSKVILNN 16

RESULT 18
 ID O6YBE8 PRELIMINARY; PRT; 11 AA.
 AC O6YBE8;
 DT 05-JUN-2004 (TREMBlrel. 27, Created)
 DT 05-JUN-2004 (TREMBlrel. 27, Last sequence update)
 DE 05-JUN-2004 (TREMBlrel. 27, Last annotation update)
 DE Wzx (Fragment).
 GN Name=wzx;
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OC NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MA053392, NZSPN00/319, NZSPN01/278, MA050663, MA052628, 01S009,
 and MA062610;
 RX MEDLINE=22979121; PubMed=14614062;
 RA Kong F., Gilbert G.L.;
 RT "Using cpSA-cpSB sequence polymorphisms and serotype-/group-specific
 PCR to predict 51 Streptococcus pneumoniae capsular serotypes." J. Med. Microbiol. 52:1047-1058(2003).
 RL EMBL; AY163227; AA060520.1; -.
 DR EMBL; AY163228; AA060522.1; -.
 DR EMBL; AY163229; AA060524.1; -.
 DR EMBL; AY163230; AA060526.1; -.
 DR EMBL; AY163231; AA060528.1; -.
 DR EMBL; AY163232; AA060530.1; -.

DR EMBL; AY163226; AAO60518.1; -.
 FT NON TER 1
 SQ SEQUENCE 11 AA; 1456 MW; EE6B29E7C721F1A1 CRC64;

Query Match 24.2%; Score 24; DB 2; Length 11;
 Best Local Similarity 80.0%; Pred. No. 4.8e+03;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 14 FLEHS 18
 DB 1 FLEHT 5

RESULT 19
 AAO60518 PRELIMINARY; PRT; 11 AA.

AC AAO60518; (TREMBlrel. 27, Created)
 DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
 DE Wzx (Fragment).
 GN WZX.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MA062610;
 RX MEDLINE=22979121; PubMed=14614062;
 RA Kong F., Gilbert G.L.;
 RT "Using cpBA-cpsB sequence polymorphisms and serotype-/group-specific PCR to predict 51 Streptococcus pneumoniae capsular serotypes.";
 RL J. Med. Microbiol. 52:1047-1058 (2003).
 DR EMBL; AY163226; AAO60518.1; -.
 FT NON TER 1
 SQ SEQUENCE 11 AA; 1456 MW; EE6B29E7C721F1A1 CRC64;

Query Match 24.2%; Score 24; DB 2; Length 11;
 Best Local Similarity 80.0%; Pred. No. 4.8e+03;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 14 FLEHS 18
 DB 1 FLEHT 5

RESULT 20
 AAO60520 PRELIMINARY; PRT; 11 AA.

AC AAO60520; (TREMBlrel. 27, Created)
 DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
 DE Wzx (Fragment).
 GN WZX.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MA051392;
 RX MEDLINE=22979121; PubMed=14614062;
 RA Kong F., Gilbert G.L.;
 RT "Using cpBA-cpsB sequence polymorphisms and serotype-/group-specific PCR to predict 51 Streptococcus pneumoniae capsular serotypes.";
 RL J. Med. Microbiol. 52:1047-1058 (2003).
 DR EMBL; AY163227; AAO60520.1; -.
 FT NON TER 1
 SQ SEQUENCE 11 AA; 1456 MW; EE6B29E7C721F1A1 CRC64;

Query Match 24.2%; Score 24; DB 2; Length 11;

Best Local Similarity 80.0%; Pred. No. 4.8e+03;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 14 FLEHS 18
 DB 1 FLEHT 5

RESULT 21
 AAO60522 PRELIMINARY; PRT; 11 AA.

AC AAO60522; (TREMBlrel. 27, Created)
 DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
 DE Wzx (Fragment).
 GN WZX.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NZSPN01/278;
 RX MEDLINE=22979121; PubMed=14614062;
 RA Kong F., Gilbert G.L.;
 RT "Using cpBA-cpsB sequence polymorphisms and serotype-/group-specific PCR to predict 51 Streptococcus pneumoniae capsular serotypes.";
 RL J. Med. Microbiol. 52:1047-1058 (2003).
 DR EMBL; AY163228; AAO60522.1; -.
 FT NON TER 1
 SQ SEQUENCE 11 AA; 1456 MW; EE6B29E7C721F1A1 CRC64;

Query Match 24.2%; Score 24; DB 2; Length 11;
 Best Local Similarity 80.0%; Pred. No. 4.8e+03;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 14 FLEHS 18
 DB 1 FLEHT 5

RESULT 22
 AAO60524 PRELIMINARY; PRT; 11 AA.

AC AAO60524; (TREMBlrel. 27, Created)
 DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
 DE Wzx (Fragment).
 GN WZX.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NZSPN01/278;
 RX MEDLINE=22979121; PubMed=14614062;
 RA Kong F., Gilbert G.L.;
 RT "Using cpBA-cpsB sequence polymorphisms and serotype-/group-specific PCR to predict 51 Streptococcus pneumoniae capsular serotypes.";
 RL J. Med. Microbiol. 52:1047-1058 (2003).
 DR EMBL; AY163229; AAO60524.1; -.
 FT NON TER 1
 SQ SEQUENCE 11 AA; 1456 MW; EE6B29E7C721F1A1 CRC64;

Query Match 24.2%; Score 24; DB 2; Length 11;
 Best Local Similarity 80.0%; Pred. No. 4.8e+03;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 FLEHT 5

RESULT 23

AAO60526 PRELIMINARY; PRT; 11 AA.

AC AAO60526; 02-MAR-2004 (TREMBlrel. 27, Created)

DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)

DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)

DE Wzx (Fragment).

GN WZX.

OS Streptococcus pneumoniae.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI_TaxID=1313;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MA050663;

RX MEDLINE=22979121; PubMed=14614062;

RA Kong F., Gilbert G.L.;

RT "Using cpeA-cpsB sequence polymorphisms and serotype-/group-specific PCR to predict 51 Streptococcus pneumoniae capsular serotypes.";

RL J. Med. Microbiol. 52:1047-1058 (2003).

DR EMBL; AY163230; AAO60526.1; -.

FT NON TER 1 1

SQ SEQUENCE 11 AA; 1456 MW; EE6B29E7C721F1A1 CRC64;

Query Match 24.2%; Score 24; DB 2; Length 11;

Best Local Similarity 80.0%; Pred. No. 4.8e+03;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 14 FLEHS 18

Db 1 FLEHT 5

RESULT 24

AAO60528 PRELIMINARY; PRT; 11 AA.

AC AAO60528; 02-MAR-2004 (TREMBlrel. 27, Created)

DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)

DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)

DE Wzx (Fragment).

GN WZX.

OS Streptococcus pneumoniae.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI_TaxID=1313;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MA052628;

RX MEDLINE=22979121; PubMed=14614062;

RA Kong F., Gilbert G.L.;

RT "Using cpeA-cpsB sequence polymorphisms and serotype-/group-specific PCR to predict 51 Streptococcus pneumoniae capsular serotypes.";

RL J. Med. Microbiol. 52:1047-1058 (2003).

DR EMBL; AY163231; AAO60528.1; -.

FT NON TER 1 1

SQ SEQUENCE 11 AA; 1456 MW; EE6B29E7C721F1A1 CRC64;

Query Match 24.2%; Score 24; DB 2; Length 11;

Best Local Similarity 80.0%; Pred. No. 4.8e+03;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 14 FLEHS 18

Db 1 FLEHT 5

RESULT 25

AAO60530

ID AAO60530 PRELIMINARY; PRT; 11 AA.

AC AAO60530; 02-MAR-2004 (TREMBlrel. 27, Created)

DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)

DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)

DE Wzx (Fragment).

GN WZX.

OS Streptococcus pneumoniae.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI_TaxID=1313;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=01S009;

RX MEDLINE=22979121; PubMed=14614062;

RA Kong F., Gilbert G.L.;

RT "Using cpeA-cpsB sequence polymorphisms and serotype-/group-specific PCR to predict 51 Streptococcus pneumoniae capsular serotypes.";

RL J. Med. Microbiol. 52:1047-1058 (2003).

DR EMBL; AY163232; AAO60530.1; -.

FT NON TER 1 1

SQ SEQUENCE 11 AA; 1456 MW; EE6B29E7C721F1A1 CRC64;

Query Match 24.2%; Score 24; DB 2; Length 11;

Best Local Similarity 80.0%; Pred. No. 4.8e+03;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 14 FLEHS 18

Db 1 FLEHT 5

RESULT 26

07TMB4 PRELIMINARY; PRT; 13 AA.

ID 07TMB4; 01-OCT-2003 (TREMBlrel. 25, Created)

AC 07TMB4; 01-OCT-2003 (TREMBlrel. 25, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE UORF4 (UORF2).

GN Name=Bf5a;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/c;

RA Jenkins Z.A., Johansson H.E.,

RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY129324; AAN17525.1; -.

DR EMBL; AY129326; AAN17530.1; -.

DR EMBL; AY129329; AAN17537.1; -.

DR GO; GO:0005737; C:cytoplasm; IDA.

DR GO; GO:0005634; C:nucleus; IDA.

DR GO; GO:0006915; P:apoptosis; IDA.

SQ SEQUENCE 13 AA; 1477 MW; 8AA04CCF18E35A5 CRC64;

Query Match 24.2%; Score 24; DB 2; Length 13;

Best Local Similarity 50.0%; Pred. No. 5.8e+03;

Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 4 GLMNN 9

Db 2 GLWSS 7

RESULT 27

09R557 PRELIMINARY; PRT; 16 AA.

ID 09R557; 01-MAY-2000 (TREMBlrel. 13, Created)

AC 09R557; 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Subtilisin GX (Fragment).
OS Bacillus sp.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1409;
RN [1]
RP SEQUENCE.
RX MEDLINE=93356814; PubMed=8352796;
RA Durham D.R.;
RT "The elastolytic properties of subtilisin GX from alkalophilic
RT Bacillus sp. strain 6644 provides a means of differentiation from
RT other subtilisins."
RL Biochem. Biophys. Res. Commun. 194:1365-1370(1993).
DR PIR: PC1299; PC1299.
SQ SEQUENCE 16 AA; 1804 MW; 000FD22FE3940C94 CRC64;
Query Match 24.2%; Score 24; DB 2; Length 16;
Best Local Similarity 42.9%; Pred. No. 7.3e+03;
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 6 WNNNOTQL 12
DB 5 WGNRRVQ 11
RESULT 28
ID 066737 PRELIMINARY; PRT; 16 AA.
AC 066737;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Transmembrane protein (Fragment).
OS Equine infectious anemia virus.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11665;
RN [1]
RP SEQUENCE FROM N.A.
RC STPAIN=Growing wild type non-cell culture adapted; TISSUE=Blood;
RX MEDLINE=92292307; PubMed=118398;
RA Peiray S.T., Flaherty M.T., Kelley M.J., Clabough D.L., Tronick S.R.,
RA Cogging L., Whetter L., Lengel C.R., Fuller F.J.;
RT "The surface envelope protein gene region of equine infectious anemia
RT virus is not an important determinant of tropism in vitro."
RL J. Virol. 66:4085-4097(1992).
DR EMBL; M87589; AAA43031.1;
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR001361; GP90_EIAV.
DR Pfam; PF00971; EIAV_GP90_1.
KW Transmembrane.
FT NON_TER 16
SQ SEQUENCE 16 AA; 1918 MW; BAE09AF8B99A3DB4 CRC64;
Query Match 24.2%; Score 24; DB 2; Length 16;
Best Local Similarity 66.7%; Pred. No. 7.3e+03;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 8 NNNOTQL 13
DB 9 NNKTEL 14
RESULT 29
ID 081VK5 PRELIMINARY; PRT; 17 AA.
AC 081VK5;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Steerin2 protein (Fragment).
GN Name=STEERIN2;
OS

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Peeters P.J., Verhaaselt P., Moechara D.W., Luyten W.H.M.L.,
RA Geyzen J.J.G.H.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ488206; CAD2559.1;
FT NON_TER 17
SQ SEQUENCE 17 AA; 2169 MW; 1F508A5824873F9E CRC64;
Query Match 24.2%; Score 24; DB 2; Length 17;
Best Local Similarity 40.0%; Pred. No. 7.8e+03;
Matches 4; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 6 WNNNOTQL 15
DB 8 WEQNTTWKL 17
RESULT 30
ID 078381 PRELIMINARY; PRT; 17 AA.
AC 078381;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Immunodeficiency virus type 1, viral sample FLPBR5F (Florida patient
DE B), partial env cds, V5 region. (Fragment).
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92271245; PubMed=1589796;
RA Ou C.-Y., Ciesielski C.A., Myers G., Banda C.I., Luo C.C.,
RA Kober B.T.M., Mullins J.I., Schochetman G., Berklman R.L.,
RA Economou A.N., Witte J.J., Furman L.J., Satten G.A., Curran J.W.,
RA Jaffe H.W.;
RT "Molecular epidemiology of HIV transmission in a dental practice."
RL Science 256:1165-1171(1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Zhang L.Q., Leigh-Brown A.J.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M92126; AAA44496.1;
DR GO; GO:0019031; C:Viral envelope; IEA.
DR InterPro; IPR000777; GP120.
FT NON_TER 1
FT NON_TER 17
SQ SEQUENCE 17 AA; 1708 MW; 347570D2D12CA370 CRC64;
Query Match 24.2%; Score 24; DB 2; Length 17;
Best Local Similarity 50.0%; Pred. No. 7.8e+03;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 7 NNNNOTQL 14
DB 4 NTNNTET 11
RESULT 31
ID 09HBD6 PRELIMINARY; PRT; 18 AA.
AC 09HBD6;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE ATP7B (Fragment).
GN Name=ATP7B;
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wang N., Wu Z.Y., Murong S.X., Lin M.T., Fang L.,
 RT "Hot point mutations of Wilson disease gene in Chinese with DNA
 sequencing."
 RL Chung-Hua Shen Ching Ko Tsa Chih 31:20-23(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Wu Z.Y., Wang N., Murong S.X.,
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF254565; AAC27542.1; -.
 FT NON_TER 1 1
 FT NON_TER 18 18
 SQ SEQUENCE 18 AA; 2100 MW; B9E7FA2BF70B0B78 CRC64;
 Query Match 24.2%; Score 24; DB 2; Length 18;
 Best Local Similarity 42.9%; Pred. No. 8.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 Oy 4 GLMNNOTQFLFH 17
 Db 13 GLW-----LEH 18
 RESULT 32
 Q9ZG65 PRELIMINARY; PRT; 18 AA.
 AC Q9ZG65;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
 DE Orotidine-5'-phosphate decarboxylase (Fragment).
 GN Name=pyrF;
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=L2 434B;
 RA Wang L., Steenburg S.D., Zheng Y., Larsen S.H.,
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF087291; AAD04068.1; -.
 FT NON_TER 1 1
 FT NON_TER 18 18
 SQ SEQUENCE 18 AA; 2026 MW; CB911767583AFAE3 CRC64;
 Query Match 24.2%; Score 24; DB 2; Length 18;
 Best Local Similarity 36.4%; Pred. No. 8.3e+03;
 Matches 4; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 Oy 1 SSDGLMNNOT 11
 Db 2 SSETSWNTRNS 12
 RESULT 33
 Q9QUY4 PRELIMINARY; PRT; 19 AA.
 AC Q9QUY4;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE Oligoendocyt-e-specific UDP-galactose:ceramide galactosyltransferase
 (Fragment).
 OS Rattus sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10118;
 RN [1]
 RP SEQUENCE.

RX MEDLINE=96085162; PubMed=8521863;
 RA Schulte S., Stoffel W.;
 RT "UDP galactose:ceramide galactosyltransferase and glutamate/aspartate
 transporter. Copurification, separation and characterization of the
 two glycoproteins."
 RL Eur. J. Biochem. 233:947-953(1995).
 SQ SEQUENCE 19 AA; 1995 MW; 0FD4BAE303B99454 CRC64;
 Query Match 24.2%; Score 24; DB 2; Length 19;
 Best Local Similarity 80.0%; Pred. No. 8.8e+03;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 2 SDGLW 6
 Db 5 STGLW 9
 RESULT 34
 Q9QV32 PRELIMINARY; PRT; 19 AA.
 AC Q9QV32;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE G3.5 ANTIGEN=ALPHA-actinin (Fragment).
 OS Rattus sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10118;
 RN [1]
 RP SEQUENCE.
 RA Price K.A., Malhotra S.K., Koke J.R.;
 RT "Localization and characterization of an intermediate filament-
 associated protein."
 RL Submitted (XXX-1993) to the EMBL/GenBank/DBJ databases.
 FT NON_TER 1 1
 FT NON_TER 19 19
 SQ SEQUENCE 19 AA; 2205 MW; 506724CB038EB879 CRC64;
 Query Match 24.2%; Score 24; DB 2; Length 19;
 Best Local Similarity 40.0%; Pred. No. 8.8e+03;
 Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 Oy 8 NNQOTQFLFH 17
 Db 7 NLAWQIFVKH 16
 RESULT 35
 LPH3 HUMAN STANDARD; PRT; 20 AA.
 ID LPH3 HUMAN
 AC P56643;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Lysophospholipase HL-60 peak 3 (EC 3.1.1.5) (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Leukemia;
 RX MEDLINE=9111958; PubMed=1471998;
 RA Garsetti D., Holtsberg F., Steiner M.R., Egan R.W., Clark M.A.;
 RT "Butyric acid-induced differentiation of HL-60 cells increases the
 expression of a single lysophospholipase."
 RL Biochem. J. 288:831-837(1992).
 CC -1- FUNCTION: Degradation of lysophosphatidylcholine.
 CC -1- CATALYTIC ACTIVITY: 2-lysophosphatidylcholine + H(2)O =
 CC glycerophosphocholine + a fatty acid anion.
 CC -1- PTM: The N-terminus is blocked.
 DR PIR; S27351; S27351.

DR GO:0004622; F:lyso-phospholipase activity; IDA.
 DR GO:0009395; P:phospholipid catabolism; NAS.
 KW Direct protein sequencing; Hydrolase; Lipid degradation.
 FT NON_TER 1
 FT SEQUENCE 20 AA; 2263 MW; 4D2986D3F21F87F CRC64;
 SQ SEQUENCE 20 AA; 2263 MW; 4D2986D3F21F87F CRC64;
 Query Match 24.2%; Score 24; DB 1; Length 20;
 Best Local Similarity 35.7%; Pred. No. 9.3e+03;
 Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 1 SSDGLMNNQTLF 14
 :|||:
 Db 2 ASISLYNNTSLSYF 15

RESULT 36
 OMPH_HAEGA STANDARD; PRT; 20 AA.
 ID OMPH_HAEGA
 AC P80451;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Major outer membrane protein (MOMP) (Outer membrane protein H) (Fragment).
 DE (Fragment).
 GN Name:ompH;
 OS Haemophilus gallinarum;
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=728;
 RN [1]
 RP SEQUENCE, AND SUBCELLULAR LOCATION.
 RC STRAIN=IDH 1645 / Serovar A1;
 RA MEDLINE=96434400; PubMed=8637367;
 RX Hartmann L., Schroeder W., Luebke-Becker A.;
 RT "A comparative study of the major outer membrane proteins of the avian
 RT Haemophilus and Pasteurella gallinarum."
 RL Zentralbl. Bakteriol. 284:47-51 (1996).
 CC -1- FUNCTION: Structural rigidity of the outer membrane of elementary
 CC bodies and porin forming, permitting diffusion of solutes through
 CC the intercellular reticulate body membrane.
 CC -1- SUBUNIT: Disulfide bond interactions within and between MOMP
 CC molecules and other components form high molecular-weight
 CC oligomers.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
 CC -1- SIMILARITY: Belongs to the Gram-negative porin family.
 KW Direct protein sequencing; Outer membrane; Porin; Transmembrane.
 FT NON_TER 20
 FT SEQUENCE 20 AA; 2142 MW; 2B0C2CED4FE635C8 CRC64;
 Query Match 24.2%; Score 24; DB 1; Length 20;
 Best Local Similarity 44.4%; Pred. No. 9.3e+03;
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 5 LMNNQOTL 13
 :|||:
 Db 3 VYNNEGTOV 11

RESULT 37
 O51950 PRELIMINARY; PRT; 20 AA.
 ID O51950
 AC O51950;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Major outer membrane protein (Fragment).
 DE Chlamydia trachomatis.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=98055742; PubMed=9395364;

RA Bobo L.D., Novak N., Munoz B., Hsieh Y.H., Quinn T.C., West S.;
 RT "Severe disease in children with trachoma is associated with
 RT persistent Chlamydia trachomatis infection."
 RL J. Infect. Dis. 176:1524-1530 (1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Bobo L., Novak N.G.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF015548; AAB95376.1; -.
 FT NON_TER 1
 FT SEQUENCE 20 AA; 2209 MW; D838001675219589 CRC64;
 SQ SEQUENCE 20 AA; 2209 MW; D838001675219589 CRC64;
 Query Match 24.2%; Score 24; DB 2; Length 20;
 Best Local Similarity 57.1%; Pred. No. 9.3e+03;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 7 NNQOTL 13
 :|||:
 Db 2 NENQTV 8

RESULT 38
 O51952 PRELIMINARY; PRT; 20 AA.
 ID O51952
 AC O51952;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Outer membrane protein (Fragment).
 DE Chlamydia trachomatis.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bobo L., Novak N.G.;
 RT "Severe disease in children with trachoma is associated with
 RT persistent Chlamydia trachomatis infection."
 RL J. Infect. Dis. 0:0-0 (1998).
 DR EMBL: AF015550; AAB95378.1; -.
 FT NON_TER 1
 FT SEQUENCE 20 AA; 2223 MW; D83804475219589 CRC64;
 SQ SEQUENCE 20 AA; 2223 MW; D83804475219589 CRC64;
 Query Match 24.2%; Score 24; DB 2; Length 20;
 Best Local Similarity 57.1%; Pred. No. 9.3e+03;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 7 NNQOTL 13
 :|||:
 Db 2 NENQTV 8

RESULT 39
 O7TMA1 PRELIMINARY; PRT; 20 AA.
 ID O7TMA1
 AC O7TMA1;
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE UORP3 (UORF1).
 DE UORP3 (UORF1).
 GN Name=Bif5a;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c;
 RA Jenkins Z.A., Johansson H.E.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY129324; AANI7524.1; -.
 DR EMBL: AY129326; AANI7529.1; -.

DR EMBL; AY129329; AAN17536.1; -.
 DR GO; GO:0005737; C:cytoplasm; IDA.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR GO; GO:0006915; P:apoptosis; IDA.
 SQ SEQUENCE 20 AA; 2265 MW; E3C2EBF80E00F3F CRC64;

Query Match 24.2%; Score 24; DB 2; Length 20;
 Best Local Similarity 50.0%; Pred. No. 9.3e+03;
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 GLMNN 9
 Db 9 GLWSSS 14

RESULT 40

090X92 PRELIMINARY; PRT; 20 AA.
 AC 090X92;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Tyrosinase (Fragment).
 GN Name=TYR;
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Deng X., Yang Y., Liu W.;
 RL Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF416915; AAL14561.1; -.
 FT NON_TER 1 1
 FT NON_TER 20 20
 SQ SEQUENCE 20 AA; 2300 MW; AD237F25FA1BA696 CRC64;

Query Match 24.2%; Score 24; DB 2; Length 20;
 Best Local Similarity 44.4%; Pred. No. 9.3e+03;
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 LMNNNOIOL 13
 Db 11 VMANTOSLL 19

Search completed: December 30, 2004, 16:31:25
 Job time : 192 secs